ICASE

DISCRETE APPROXIMATION METHODS

FOR PARAMETER IDENTIFICATION IN DELAY SYSTEMS

I. Gary Rosen

Report No. 81-36

November 9, 1981

INSTITUTE FOR COMPUTER APPLICATIONS IN SCIENCE AND ENGINEERING

NASA Langley Research Center, Hampton, Virginia 23665

Operated by the

UNIVERSITIES SPACE USRA RESEARCH ASSOCIATION

DISCRETE APPROXIMATION METHODS FOR PARAMETER IDENTIFICATION IN DELAY SYSTEMS

I. G. Rosen

Department of Mathematics, Bowdoin College
Brunswick, MAINE 04011

ABSTRACT

We construct approximation schemes for parameter identification problems in which the governing state equation is a linear functional differential equation of retarded type. The basis of the schemes is the replacement of the parameter identification problem having an infinite dimensional state equation by a sequence of approximating parameter identification problems in which the states are given by finite dimensional discrete difference equations. The difference equations are constructed using linear semigroup theory and rational function approximations to the exponential. Sufficient conditions are given for the convergence of solutions to the approximating problems, which can be obtained using conventional methods, to solutions to the original parameter identification problem. Finite difference and spline based schemes using Padé rational function approximations to the exponential are constructed and shown to satisfy the sufficient conditions for convergence. A discussion and analysis of numerical results obtained through the application of the schemes to several examples is included.

Work supported in part by the Air Force Office of Scientific Research under Contract AFOSR 76-3092D, in part by the National Science Foundation under Grant NSF-MCS 7905774-02, and in part by the U. S. Army Research Office under Contract ARO-DAAG29-79-C-0161. Additional research was carried out while the author was in residence at the Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, VA under NASA Contract Nos. NAS1-15810 and NAS1-16394.

1. Introduction

The purpose of this paper is the investigation of approximation methods for the identification of parameters in control systems where the state equation is a linear retarded functional differential equation (LRFDE). The parameters which we are interested in being able to estimate include system coefficients, initial conditions and the delays themselves. The methods which we shall discuss are based upon the discrete approximation framework for the integration of LRFDE initial value problems developed in [27] and [28]. The approach we take is to first replace the LRFDE which governs the dynamics of the system by an equivalent abstract evolution equation set in an infinite dimensional Hilbert space. The abstract evolution equation is then approximated by a finite dimensional discrete difference equation. This in turn leads to a totally discrete finite dimensional approximating parameter identification problem which can then be solved through the use of standard techniques (see for instance [29]) and software packages which are readily available. That the solutions to these approximating problems in some sense approximate solutions to the original parameter identification problem is the primary result discussed in this paper.

As is pointed out in [7] very little regarding the parameter identification problem for delay systems (PIDDS) appears in the literature. This is especially true for the case in which the delays are among the parameters to be identified. More recently, however, research in this area has been undertaken. Banks, Burns and Cliff [7] have extended the approximation framework which had previously been developed to solve optimal control problems governed by delay differential systems [1], [4], [5], [9] [12], [17], [22] so as to be applicable to the PIDDS as well. Their approach is based upon semi-discrete methods (i.e. those methods in which the LRFDE state equation is replaced by an approximating ordinary differential equation) for the integration of the delay differential equation which governs the dynamics of the system. The convergence arguments given by the authors rely heavily upon an abstract formulation of the problem which permits the use of linear semigroup theory and the associated approximation results which have been developed for application in such a setting. Their approximation framework is applicable to an extremely wide class of problems, includes methods having an arbitrarily high order of convergence and is capable of identifying the delays which appear in the state equation.

Banks in [2] develops spline based semi-discrete approximation schemes which are applicable to PIDDS in which the delays are not among the parameters to be estimated and in which the state equation is a nonlinear delay system satisfying global Lipschitz and differentiability conditions. While an equivalent abstract formulation of the problem is employed, in this treatment, the author has avoided the use of semigroup theory entirely. Instead, the convergence of the approximation schemes is argued via the dissipativeness of the nonlinear operators involved and the Gronwall inequality. In [8] and [14] the ideas discussed in [2] are further extended so as to be applicable to problems in which the delays are to be identified as well.

In addition to the construction of approximation schemes, a discussion of modeling problems arising in physiology, enzyme kinetics and unsteady aerodynamics which involve parameter identification and control for delay systems can be found in [3] and [7].

The framework which we shall develop is closely related to the ideas contained in [7]. We treat essentially the same class of problems, rely upon the same abstract formulation, apply many of the same functional analytic techniques to argue convergence and in fact, incorporate the same state discretizations as those discussed in [7]. The primary difference in the two approaches, however, is that our methods result in a complete discretization of the problem and hence require no further approximation when implemented. The methods included in our framework are capable of identifying delays and the integration schemes which they rely upon may be constructed with an arbitrarily high order of convergence. A detailed comparison of the performance of the semi-discrete schemes developed in [7] and totally discrete schemes similar to those which will be discussed here when applied to parameter identification problems in which the delays themselves are not among the parameters to be estimated can be found in [11].

An alternative treatment of the problem of approximating solutions to the PIDDS which is also based upon discrete approximation via difference equations is given by Burns and Hirsch in [13]. These authors have taken a somewhat more straightforward approach by studying a specific scheme (as opposed to developing an approximation framework as is done in [7] and will be done here) which can be applied to PIDDS in which the LRFDE contains a single discrete delay term only. (The schemes developed here and those discussed in [7] are capable of handling equations which contain multiple discrete delay terms as well as a distributed delay term). The approximating difference equation is derived via the modification of standard numerical integration schemes for ordinary differential equations (i.e. Euler's method, fourth order Runge, Kutta, etc.) so as to be applicable to delay differential systems. The authors are able to argue first order convergence for the Euler based scheme directly and hence can avoid the necessity for a functional analytic formulation of the problem. Computational evidence supporting the feasibility of extending these ideas to higher order schemes is also provided. However, the authors point out that the possibility of extending the relevant convergence arguments is uncertain. The Burns and Hirsch paper also addresses the difficulties which can arise in the construction of approximation schemes for PIDDS in which the delays are to be identified due to the fact that solutions to delay differential systems may not be smooth with respect to the delays. This can pose problems since most standard optimization packages require differentiability with respect to the parameters.

Although it does not concern itself with the PIDDS directly, the work of Banks and Kunisch [10] should also be included in this historical outline. In this paper the authors treat parameter identification problems in which the governing state equations are semi-linear parabolic or hyperbolic partial differential equations. The approach that they take is similar to the one that is taken in [7]. Indeed, the infinite dimensional identification problem is replaced by an equivalent abstract formulation which is then used to develop finite dimensional semi-discrete approximation schemes. In a similar manner, the totally discrete schemes which will be developed below could easily be modified so as to be applicable to parameter identification problems with partial differential state equations.

We conclude this section with a brief outline of our presentation. In Section 2 we state the PIDDS and show that it can be reformulated as an equivalent parameter identification problem in which the state equation is an abstract evolution equation set in an infinite dimensional Hilbert space. In Section 3 we establish approximation results for abstract evolution equations while in Section 4 we use these results to construct the approximating parameter identification problems and to show that under the appropriate hypotheses, solutions to the approximating problems converge to solutions to the PIDDS. In Section 5 we construct actual approximation schemes which satisfy the hypotheses and conditions necessary for convergence, while in Section 6 we discuss and analyze numerical results obtained through the application of these schemes to several examples.

The notation we employ is, for the most part, standard. The symbol $\mathcal{L}_{n\times n}$ is used to denote the space of n square matrices. We denote the space of functions defined on (a,b) with range in R^n and p continuous derivatives by $C_p^n(a,b)$. The space of piecewise continuous functions and the space of continuous functions on (a,b) with range in R^n are denoted by $PC^n(a,b)$ and $C^n(a,b)$ respectively. The Lebesgue spaces of R^n valued functions on (a,b) are denoted by $L_p^n(a,b)$ while the Sobolev spaces of functions ϕ with $\phi^{(m-1)}$ absolutely continuous and $\phi^{(m)}$ in $L_p^n(a,b)$ are denoted by $W_{m,p}^n(a,b)$. For a function $\phi \in W_{1,2}^n(a,b)$ we shall use the notations $D\phi$ and $\dot{\phi}$ interchangeably to denote the derivative of ϕ. Finally for a linear operator $\mathcal{T}$, the symbols $D(\mathcal{T})$ and $R(\mathcal{T})$ are used to denote the domain of $\mathcal{T}$ and the range of $\mathcal{T}$ respectively.

2. The PIDDS and its Abstract Formulation

In this section we formulate the parameter identification problem for delay systems and show that it has an equivalent formulation, whereby the dynamics of the governing control system in the form of an LRFDE are replaced by an abstract evolution equation set in an infinite dimensional Hilbert space. Since the PIDDS and the associated approximation schemes which we shall develop are closely related to the problem and schemes discussed by Banks, Burns and Cliff [7] the reader is instructed to note the similarities which exist between the material and notation to follow in this section and that which is contained in sections 2 and 2.1 of [7].

We begin with the definition of the admissible initial data/parameter set and a formal statement of the PIDDS. Let $r > 0$ and Ω a compact convex subset of R^{μ} be given. Define the compact convex set $Q \subset R^{\mu+\nu}$ by $Q \equiv \Omega \times H$ where

$$H \equiv \{h = (r_1, r_2 \ldots r_\nu \in R^\nu \mid 0 \leq r_i \leq r_{i+1} \leq r, \quad i = 1, 2 \ldots \nu - 1\} \ .$$

In addition let S be a compact convex subset of $R^n \times L_2^n(-r,0)$ and define

$$\Gamma = S \times Q = S \times \Omega \times H$$

to be the admissible initial data/parameter set. We further assume that we have been provided with an input/output pair $(u,\zeta) \in PC^m(0,T) \times C^\ell(0,T)$ for some $T > 0$. We refer to (u,ζ) as an input/output pair since it is assumed that if given input $u \in PC^m(0,T)$ the physical system to be identified produces output $\zeta \in C^\ell(0,T)$.

With the above definitions in hand, we can state the PIDDS:

(PIDDS): Given an input/output pair $(u,\zeta) \in PC^m(0,T) \times C^\ell(0,T)$ for some $T > 0$, find $\gamma^* = (\eta^*, \phi^*, q^*) = (\eta^*, \phi^*, \alpha^*, h^*) \in \Gamma$ which minimizes

$$(2.1) \qquad J(\gamma) = |y(0;\gamma,u) - \zeta(0)|^2_{w_1} + |y(T;\gamma,u) - \zeta(T)|^2_{w_2} + \int_0^T |y(t;\gamma,u) - \zeta(t)|^2_{w_3}$$

subject to

$$(2.2)\qquad \dot{x}(t) = L(q)x_t + B(\alpha)u(t) \qquad t \in [0,T]$$

$$(2.3)\qquad x(0) = \eta \qquad x_0 = \phi$$

$$(2.4)\qquad y(t) = C(\alpha)\,x(t) + D(\alpha)\,u(t)$$

where for each $\alpha \in \Omega$, $B(\alpha)$, $C(\alpha)$ and $D(\alpha)$ are $n\times m$, $\ell\times n$. and $\ell\times m$ matrices respectively, $|\cdot|_{w_j}$ $j = 1,2,3$ represent appropriately weighted (application dependent) norms on R^{ℓ}, x_t denotes the function $\theta \to x(t+\theta) - r \leq \theta \leq 0$ and the notation $y(\cdot\,;\gamma, u)$ is employed in order to exhibit the explicit dependence of the output y of the theoretical system on the initial conditions and parameter values γ and the given input u. For each $q = (\alpha,h) = (\alpha,r_1,r_2\ .\ r_\nu) \in Q$ the operator $L(q)\colon L_2^n(-r,0) \to R^n$ is assumed to be of the form

$$L(q)\phi = \sum_{i=0}^{\nu} A_i(\alpha)\phi(-r_i) + \int_{-r_\nu}^{0} K(\alpha,\theta)\,\phi(\theta)d\theta$$

with $r_0 \equiv 0$ and where for each $\alpha \in \Omega$ $A_i(\alpha)$ $i = 0,1,2,\ldots,\nu$ are $n\times n$ matrices and $\theta \to K(\alpha,\theta)$ is an $n\times n$ matrix valued function in $L_2((-r,0), L_{n\times n})$. It is assumed that $A_i(\alpha)$ $i = 0,1,2,\ldots,\nu$, $B(\alpha)$, $C(\alpha)$, $D(\alpha)$, $K(\alpha,\cdot)$ are continuous in α.

Before we go on to discuss the parameter identification problem, let us take a moment to consider the LRFDE initial value problem given by (2.2) (2.3). Given $\gamma = (\eta,\phi,q) \in \Gamma$, a solution to the initial value problem is a function $x\colon[-r,T] \to R^n$ such that $x \in W_{1,2}^n(0,T)$, x satisfies equation (2.2) almost everywhere on $[0,T]$, $x(0) = \eta$, $x_0 = \phi$. Standard arguments [24] can be used to demonstrate that for each $\gamma \in \Gamma$ (2.2) (2.3) has a unique solution which depends continuously upon $\gamma \in \Gamma$ and the non-homogeneous term u (as an element of $L_2^m(0,T)$.). The notation $x(t;\gamma,u)$ (and $x_t(\gamma,u)$) will be used to denote this unique solution (and its past history on $[t-r,t]$) to (2.2) (2.3) corresponding to a particular choice of $\gamma \in \Gamma$ and $u \in L_2^m(0,T)$.

Remark: One might be tempted to question the validity of choosing a least squares payoff functional of the form given in (2.1) for the PIDDS since in actual practice it is usually the case that for a given input u, output can only be measured at discrete times $0 \leq t_0 < t_1 \ldots < t_m \leq T$. In this instance a more appropriate choice for a payoff functional would be the one used in [7] which is given by

$$J(\gamma) = \frac{1}{2}\sum_{j=0}^{m}|y(t_j;\gamma,u) - \zeta_j|^2$$

where the $\{\zeta_j\}_{j=0}^{m}$ are the given discrete output observations obtained from the actual system which is to be identified. Oddly enough, it is the discrete nature of the approximation schemes to be discussed which necessitates the use of the distributed payoff functional given by (2.1). However, this restriction can be circumvented via the use of an interpolation scheme applied either to the observational data provided in order to generate a continuous observation $\hat{\zeta}(\cdot) \in C^{\ell}(0,T)$ or to the discrete output generated by the difference equation based approximation schemes. The latter approach is the one which is employed in [13] in order to overcome this very same problem.

We next show that the PIDDS has an equivalent formulation as a parameter identification problem in which the governing state equation is given by an abstract evolution equation set in the Hilbert space Z given by

$$Z \equiv R^n \times L_2^n(-r,0) \ .$$

with inner product

$$\langle \cdot,\cdot \rangle_Z = \langle \cdot,\cdot \rangle_{R^n} + \langle \cdot,\cdot \rangle_{L_2} \ .$$

The quantity r which appears in the definition of the space Z is as had been defined previously. For $q = (\alpha,h) \in Q$ and $(\eta,\phi) \in Z$ we define the parameterized family of operators $S(t;q): Z \to Z$ for $t \geq 0$ by

$$S(t;q)(\eta,\phi) = (x(t;(\eta,\phi,q),0),\ x_t((\eta,\phi,q),0))$$

where $x(\cdot,(\eta,\phi,q),0)$ denotes the unique solution to (2.2) (2.3) corresponding to $q \in Q$, $(\eta,\phi) \in Z$ and $u \equiv 0$. In light of the existence, uniqueness and continuous dependence results for solutions to the initial value problem (2.2) (2.3) discussed earlier, it is not difficult to show that for each $q \in Q$ the operators $\{S(t;q): t \geq 0\}$ form a C_0 semigroup of bounded linear operators on Z . Furthermore, for each $q \in Q$ the infinitesimal generator $\mathcal{A}(q): D(\mathcal{A}(q)) \subset Z \to Z$ of the semigroup and its domain of definition (which is independent of q) can be calculated. They are given by

$$D(\mathcal{A}(q)) = D = \{(\eta,\phi) \in Z \,|\, \phi \in W^n_{1,2}(-r,0), \eta = \phi(0)\}$$

$$\mathcal{A}(q)(\phi(0),\phi) = (L(q)\phi, D\phi) \ .$$

Turning our attention next to the non-homogeneous equation, for each $\alpha \in \Omega$ we define the operator $\hat{B}(\alpha): R^m \to Z$ by $\hat{B}(\alpha)u = (B(\alpha)u,0)$ and consider

$$z(t;\gamma,u) = S(t;q)(\eta,\phi) + \int_0^t S(t-\sigma;q)\hat{B}(\alpha)u(\sigma)\,d\sigma \tag{2.5}$$

$0 \leq t \leq T$ for each $\gamma = (\eta,\phi,q) = (\eta,\phi,\alpha,h) \in \Gamma$ and $u \in L_2^m(0,T)$. Using standard results from linear semigroup theory [20] it is easily verified that the expression for z given in (2.5) is well defined and continuous in t. Furthermore, under the somewhat more restrictive conditions that $(\eta,\phi) \in D$ and $u \in C_1^m(0,T)$ we have that (2.5) is the unique strong solution to the initial value problem in Z given by

$$\dot{z}(t) = \mathcal{A}(q)z(t) + \hat{B}(\alpha)u(t) \tag{2.6}$$

$$z(0) = (\eta,\phi) \tag{2.7}$$

It can be shown (see [4], [5]) that under the same conditions

$$w(t) \equiv (x(t;\gamma,u)\ ,\ x_t(\gamma,u))$$

is a strong solution to the initial value problem given by (2.6) (2.7) as well. It therefore must follow that $z(t)$ and $w(t)$ coincide for $0 \leq t \leq T$. By making use of standard density and continuous dependence arguments the equivalence of z and w can be extended so as to hold for all $(\eta,\phi) \in Z$ and $u \in L_2^m(0,T)$. We state this conclusion in the form of a theorem.

<u>Theorem</u> 2.1 Let $x(\cdot;\gamma,u)$ denote the unique solution to the LRFDE initial value problem (2.2) (2.3) corresponding to $\gamma \in \Gamma$ and $u \in L_2^m(0,T)$. Then for $0 \leq t \leq T$ we have

$$z(t;\gamma,u) = (x(t;\gamma,u),\ x_t(\gamma,u))\ .$$

In the light of Theorem 2.1 above, the equivalence which exists between solutions to the LRFDE initial value problem (2.2) (2.3) and the Z valued function given by expression (2.5) permits the reformulation of the PIDDS as an equivalent (1-1 correspondence between solutions) parameter identification problem in which the governing state equation is now given by (2.5). Indeed, if we define for each $\alpha \in \Omega$ the operator $\hat{C}(\alpha): Z \to R^{\ell}$ by $\hat{C}(\alpha)(\eta,\phi) = C(\alpha)\eta$ then the PIDDS is equivalent to the following abstract parameter identification problem.

(APIDDS): Given an input/output pair $(u,\zeta) \in PC^m(0,T) \times C^{\ell}(0,T)$ for some $T > 0$, find $\gamma^* = (\eta^*,\phi^*,q^*) = (\eta^*,\phi^*,\alpha^*,h^*) \in \Gamma$ which minimizes $J(\gamma)$ given by (2.1) subject to

$$z(t) = S(t;q)(\eta,\phi) + \int_0^t S(t-\sigma;q)\hat{B}(\alpha)u(\sigma)d\sigma \tag{2.8}$$

$$y(t) = \hat{C}(\alpha)z(t) + D(\alpha)u(t) \qquad t \in [0,T] . \tag{2.9}$$

The fact that there exists a 1-1 correspondence between solutions to the APIDDS above and solutions to the PIDDS forms the basis for the approximation schemes which we construct in the succeeding sections. Indeed, we shall obtain approximate solutions to the PIDDS by constructing convergent approximation schemes which are applicable to the APIDDS and which are based upon finite dimensional difference equation approximation of the state equation given by (2.8).

3. Approximation Results for the Abstract State Equation

Fundamental to our approximation schemes for the PIDDS is the construction of convergent finite dimensional discrete approximation schemes for the state equation given by (2.8). Indeed, the approach we take is to replace the APIDDS by a parameter identification problem in which the governing state equation is now a finite dimensional linear non-homogeneous difference equation which depends upon $q \in Q$. For each $q \in Q$, the solutions to these difference equations will, in some sense, approximate the solutions of (2.8). If we let N represent the degree of approximation, and if the N^{th} approximating problem is solved for γ_N^*, an element in the admissible initial data/parameter set Γ which minimizes a discrete least squares payoff functional which approximates (2.1), then we shall see that the compactness of Γ and the convergence of the state approximation are sufficient to guarantee the existence of a subsequence $\{\gamma_{N_k}^*\}$ of $\{\gamma_N^*\}$ and a $\gamma^* \in \Gamma$ such that $\gamma_{N_k}^* \to \gamma^*$ as $k \to \infty$ with γ^* a solution to the APIDDS. In the light of the equivalence established in Section 2, it then must necessarily follow that γ^* is a solution to the PIDDS as well.

The above ideas will be made precise in Section 4. In this section, however, we shall concentrate exclusively on the abstract approximation results which are fundamental to the construction of convergent approximations to the state equation. These approximation results can be considered to be discrete analogs of the well known Trotter-Kato Theorem [20] which is frequently employed to establish the convergence of semi-discrete approximations to semigroups of operators. Although the theorems we shall state and prove below are quite similar to results appearing in [28], a direct application of the treatment in [28] to establish the convergence of the state approximations for the PIDDS is not possible. This is a consequence of the fact that since the delays can appear among the unknown parameters, the largest delay, r_ν, which plays a crucial role in our formulation, may no longer remain fixed with respect to the degree of approximation.

In what is to follow we shall employ the following conventions. For a rational function of a complex variable $r(z) = p(z)/q(z)$ we denote the degree of $r(z) =$ degree of $p(z) -$ degree of $q(z)$ by $\deg r(z)$. For $\mathcal{T}: D(\mathcal{T}) \subset H \to H$ a linear transformation on a Hilbert space H we say $\mathcal{T} \in G(M,\beta)$ if $\mathcal{T}$ is the infinitesimal generator of a C_0 semigroup of operators $\{T(t): t \geq 0\}$ satisfying $|T(t)| \leq Me^{\beta t}$. Furthermore, for $\lambda \in \rho(\mathcal{T})$, the resolvent set of $\mathcal{T}$, we denote the resolvent of $\mathcal{T}$, $(\lambda I - \mathcal{T})^{-1}$ by $R_\lambda(\mathcal{T})$.

For $\mathcal{T} \in G(M,\beta)$ and $\lambda \in \rho(\mathcal{T})$, using standard results from the theory of semigroups of operators (see [28]) the following two bounds which are employed frequently below can be established:

(3.1) $$|R_\lambda(\mathcal{T})| \leq \frac{M}{\mathrm{Re}\lambda - \beta}$$

(3.2) $$|\mathcal{T}R_\lambda(\mathcal{T})| \leq 1 + \frac{|\lambda|M}{\mathrm{Re}\lambda - \beta} \equiv M_\lambda .$$

We formulate our approximation framework in the same general setting as the one used in [7]. Let $(Z,\langle\cdot,\cdot\rangle)$, $(Z_N,\langle\cdot,\cdot\rangle_N)$ $N = 1,2\ldots$ be Hilbert spaces with norms $|\cdot|$ and $|\cdot|_N$ respectively. For each $N = 1,2\ldots$ let X_N be a closed subspace of Z_N and let $\Pi_N: Z_N \to X_N$ be the orthogonal projection of Z_N onto X_N with respect to the $\langle\cdot,\cdot\rangle_N$ inner product. Let $\mathcal{I}_N: Z \to Z_N$ be a mapping which is onto Z_N and which satisfies

$|\mathcal{I}_N z|_N \leq |z|$ for each $z \in Z$. Finally, we define $P_N: Z \to X_N$ by $P_N = \Pi_N \mathcal{I}_N$, and note that $|P_N z|_N \leq |z|$ for each $z \in Z$.

Our basic approximation result is given in Theorem 3.1 below. The proof of Theorem 3.1 relies heavily upon the following lemma which is due to Hersh and Kato [18]. The proof of Lemma 3.1 in the form in which it is stated here can be found in [28].

Lemma 3.1 Suppose

(1) $\mathcal{T} \in G(M,\beta)$ is the infinitesimal generator of the C_0 semigroup of operators $\{T(t): t \geq 0\}$.

(2) $r(z)$ is a rational function of the complex variable z which satisfies

(a) $|e^z - r(z)| = O(|z|^{m+1})$ $z \to 0$ with $m > 0$

(b) $\deg r(z) \leq m+1$

(c) $r(z)$ has no poles in $\{z \in C: \mathrm{Re}\, z \leq 0\}$.

Then there exist positive constants $\hat{M}, \varepsilon$ independent of $\mathcal{T} \in G(M,\beta)$ such that the operator $r(h\mathcal{T})$ exists. Moreover for $f \in D(\mathcal{T}^{m+1})$ we have

$$|T(h)f - r(h\mathcal{T})f| \leq \hat{M}e^{\beta h}|\mathcal{T}^{m+1}f|h^{m+1}$$

for all h with $0 \leq h \leq \varepsilon$.

Remark 3.1 Our concern with the existence of the operator $r(h\mathcal{T})$ is necessitated by the fact that $r(z) = p(z)/q(z)$ is a rational function. The existence of the operator $r(h\mathcal{T})$ is dependent therefore upon the existence of the operator inverse of $q(h\mathcal{T})$.

Theorem 3.1 Let Z, Z_N, X_N and P_N be as they have been defined above and let T be a fixed positive real number. Suppose for some M, β we have that $A \in G(M,\beta)$ is the infinitesimal generator of the C_0 semigroup of operators on Z, $\{S(t): t \geq 0\}$ and $A_N \in G(M,\beta)$ is the infinitesimal generator of the C_0 semigroup of operators on X_N, $\{S_N(t): t \geq 0\}$ $N = 1,2\ldots$. Suppose further that

(1) There exists $\mathcal{D} \subset D(A)$, a dense subset of Z such that $R_\lambda(A)\mathcal{D} \subset \mathcal{D}$ for each $\lambda \in C$ with $\mathrm{Re}\lambda > \beta$ and for each $z \in \mathcal{D}$ we have

$$|A_N P_N z - P_N A z|_N \to 0 \quad \text{as} \quad N \to \infty$$

(2) $c(z)$ is a rational function of the complex variable $z \in C$ such that

(a) $|c(z) - e^z| = O(|z|^{m+1})$ as $z \to 0$ with $m > 0$

(b) $\deg c(z) \leq m + 1$

(c) $c(z)$ has no poles in $\{z \in C: \mathrm{Re}\, z \leq 0\}$

(3) $\{r_\nu^N\}_{N=1}^\infty$ is a sequence of positive real numbers satisfying $0 < r_\nu^N \leq r < \infty$ $N = 1,2\ldots$

(4) $\{\rho^N\}_{N=1}^\infty$ is a sequence of positive integers determined by the following relation

$$\rho^N \frac{r_\nu^N}{N} \leq T < (\rho^N + 1)\, \frac{r_\nu^N}{N} \qquad N = 1,2\ldots$$

Then there exists an $\bar{N}$ such that the operators on X_N given by $c\left(\frac{r_\nu^N}{N} A_N\right)$ exist for all $N > \bar{N}$ and moreover if the infinite collection of operators

$$\left\{ c\left(\frac{r_\nu^N}{N} A_N\right)^k \right\}_{k=0}^{\rho^N} \qquad N > \bar{N}$$

are uniformly bounded, then given $\epsilon > 0$, there exists an $\hat{N} > \bar{N}$ such that

$$\left| c\left(\frac{r_\nu^N}{N} A_N\right)^k P_N z - P_N S\left(\frac{k r_\nu^N}{N}\right) z \right|_N < \epsilon$$

$k = 0,1,2 \ldots \rho^N$ for all $N > \hat{N}$ and each $z \in Z$. (Equivalently stated: $\left| c\left(\frac{r_\nu^N}{N} A_N\right)^k P_N z - P_N S\left(\frac{k r_\nu^N}{N}\right) z \right|_N \longrightarrow 0$ as $N \to \infty$ uniformly in k, $k \in \{0,1,2 \ldots \rho^N\}$)

Proof

The existence of an $\bar{N}$ such that the operators $c\left(\frac{r_\nu^N}{N} A_N\right)$ exist for all $N > \bar{N}$ is a consequence of Lemma 3.1. We next assume that the operators $\left\{ c\left(\frac{r_\nu^N}{N} A_N\right)^k \right\}_{k=0}^{\rho^N}$ $N > \bar{N}$ are uniformly bounded. Let M_0 be such that

$$\left| c\left(\frac{r_\nu^N}{N} A_N\right)^k \right|_N \le M_0 \tag{3.3}$$

$k = 0,1,2 \ldots \rho^N$, $N > \bar{N}$. Then for $z \in Z$ and $k = 0,1,2 \ldots \rho^N$ we have

$$\begin{aligned}
&\left| c\left(\frac{r_\nu^N}{N} A_N\right)^k P_N z - P_N S\left(\frac{k r_\nu^N}{N}\right) z \right|_N \\
&\le \left| \left[c\left(\frac{r_\nu^N}{N} A_N\right)^k P_N - S_N\left(\frac{k r_\nu^N}{N}\right) P_N \right] z \right|_N \\
&+ \left| \left[S_N\left(\frac{k r_\nu^N}{N}\right) P_N - P_N S\left(\frac{k r_\nu^N}{N}\right) \right] z \right|_N \\
&\equiv T_1^N + T_2^N .
\end{aligned}$$

If one compares the hypotheses of the present theorem with those of Theorem 3.1 in [7] it is clear that an application of the latter result implies that $T_2^N \to 0$ as $N \to \infty$ uniformly in k, $k \in \{0,1,2,\ldots\rho^N\}$.

We next consider the term T_1^N. Applying Lemma 3.1, the fact that $A_N \in G(M,\beta)$, (3.2) and (3.3), the following estimate can be made. For $\lambda \in C$ with $\operatorname{Re}\lambda > \beta$ we have

$$\left|\left[c\left(\frac{r_\nu^N}{N}A_N\right)^k - S_N\left(\frac{kr_\nu^N}{N}\right)\right]R_\lambda(A_N)^{m+1}P_Nz\right|_N$$

$$= \left|\sum_{j=0}^{k-1} c\left(\frac{r_\nu^N}{N}A_N\right)^j\left[c\left(\frac{r_\nu^N}{N}A_N\right) - S_N\left(\frac{r_\nu^N}{N}\right)\right]S_N\left((k-j-1)\frac{r_\nu^N}{N}\right)R_\lambda(A_N)^{m+1}P_Nz\right|_N$$

$$\leq M_0\sum_{j=0}^{k-1}\left|\left[c\left(\frac{r_\nu^N}{N}A_N\right) - S_N\left(\frac{r_\nu^N}{N}\right)\right]S_N\left((k-j-1)\frac{r_\nu^N}{N}\right)R_\lambda(A_N)^{m+1}P_Nz\right|_N$$

$$\leq M_0\hat{M}\sum_{j=0}^{k-1} e^{\frac{\beta r_\nu^N}{N}}\left(\frac{r_\nu^N}{N}\right)^{m+1}\left|A_N^{m+1}S_N\left((k-j-1)\frac{r_\nu^N}{N}\right)R_\lambda(A_N)^{m+1}P_Nz\right|_N$$

$$= M_0\hat{M}\sum_{j=0}^{k-1} e^{\frac{\beta r_\nu^N}{N}}\left(\frac{r_\nu^N}{N}\right)^{m+1}\left|S_N\left((k-j-1)\frac{r_\nu^N}{N}\right)A_N^{m+1}R_\lambda(A_N)^{m+1}P_Nz\right|_N$$

$$\leq M_0\hat{M}\sum_{j=0}^{k-1} e^{\frac{\beta r_\nu^N}{N}}\left(\frac{r_\nu^N}{N}\right)^{m+1}\left|S_N\left((k-j-1)\frac{r_\nu^N}{N}\right)\right|\left|(A_NR_\lambda(A_N))^{m+1}\right|\left|P_Nz\right|_N$$

$$\leq M_0M\hat{M}\sum_{j=0}^{k-1}\left(\frac{r_\nu^N}{N}\right)^{m+1} e^{\frac{\beta r_\nu^N}{N}}\,e^{\beta(k-j-1)\frac{r_\nu^N}{N}}\left|A_NR_\lambda(A_N)\right|^{m+1}|z|$$

$$\leq M_0\hat{M}MM_\lambda^{m+1}|z|\sum_{j=0}^{k-1}\left(\frac{r_\nu^N}{N}\right)^{m+1} e^{\beta(k-j)\frac{r_\nu^N}{N}}$$

$$\le M_0\hat{M}MM_\lambda^{m+1}|z|e^{\beta\rho^N\frac{r_\nu^N}{N}}\sum_{j=0}^{k-1}\left(\frac{r_\nu^N}{N}\right)^{m+1}$$

$$\le M_0\hat{M}MM_\lambda^{m+1}|z|e^{\beta T}\rho^N\left(\frac{r_\nu^N}{N}\right)\left(\frac{r_\nu^N}{N}\right)^m$$

$$\le M_0\hat{M}MM_\lambda^{m+1}T\ e^{\beta T}|z|\left(\frac{r_\nu^N}{N}\right)^m$$

$$= \alpha|z|\left(\frac{r_\nu^N}{N}\right)^m$$

where $\alpha \equiv M_0\hat{M}MM_\lambda^{m+1}T\ e^{\beta T}$.

Under the present hypotheses, it is not difficult to demonstrate (see Theorem 4.9 of [28]) that for each $z \in Z$ and each $\lambda \in C$ with $\mathrm{Re}\lambda > \beta$

$$|[P_N R_\lambda(A)^{m+1} - R_\lambda(A_N)^{m+1}P_N]z|_N \to 0 \qquad \text{as} \quad N \to \infty .$$

Therefore it follows that

$$\left|\left[c\left(\frac{r_\nu^N}{N}A_N\right)^k - S_N\left(\frac{kr_\nu^N}{N}\right)\right] P_N R_\lambda(A)^{m+1}z\right|_N$$

$$\le \left|\left[c\left(\frac{r_\nu^N}{N}A_N\right)^k - S_N\left(\frac{kr_\nu^N}{N}\right)\right]\left[P_N R_\lambda(A)^{m+1} - R_\lambda(A_N)^{m+1}P_N\right]z\right|_N$$

$$+ \left|\left[c\left(\frac{r_\nu^N}{N}A_N\right)^k - S_N\left(\frac{kr_\nu^N}{N}\right)\right]R_\lambda(A_N)^{m+1}P_N z\right|_N$$

$$\le \left(M_0 + M e^{\beta\rho^N\frac{r_\nu^N}{N}}\right)\left|\left[P_N R_\lambda(A)^{m+1} - R_\lambda(A_N)^{m+1}P_N\right]z\right|_N$$

$$+ \alpha|z|\left(\frac{r_\nu^N}{N}\right)^m$$

$$\leq (M_0 + M e^{\beta T}) \,|\, [P_N R_\lambda (A)^{m+1} - R_\lambda (A_N)^{m+1} P_N] z \,|_N$$

$$+ \alpha |z| \left(\frac{r}{N}\right)^m \to 0 \quad \text{as} \quad N \to \infty$$

and hence

$$\left| \left[c\left(\frac{r_\nu^N}{N} A_N \right)^k - S_N \left(\frac{kr_\nu^N}{N} \right) \right] P_N z \right|_N \to 0 \quad \text{as} \quad N \to \infty$$

uniformly in k, $k \in \{0,1,2\ldots\rho^N\}$ for each $z \in R(R_\lambda(A)^{m+1})$. However, $R(R_\lambda(A)^{m+1}) = D(A^{m+1})$ which is a dense subset of Z (see [23]). Furthermore the operators on Z given by $\left[c\left(\frac{r_\nu^N}{N} A_N\right)^k - S_N\left(\frac{kr_\nu^N}{N}\right) \right] P_N$ $k = 0,1,2\ldots\rho^N$ are uniformly bounded in N for all $N > \bar{N}$. Indeed

$$\left| \left[c\left(\frac{r_\nu^N}{N} A_N \right)^k - S_N \left(\frac{kr_\nu^N}{N} \right) \right] P_N \right| \leq M_0 + M e^{\beta T}$$

$k = 0,1,2\ldots\rho^N$ all $N > \bar{N}$. Therefore

$$\left| \left[c\left(\frac{r_\nu^N}{N} A_N \right)^k - S_N \left(\frac{kr_\nu^N}{N} \right) \right] P_N z \right|_N \to 0 \quad \text{as} \quad N \to \infty$$

uniformly in k, $k \in \{0,1,2\ldots\rho^N\}$ for each $z \in Z$ and the theorem is proven.

Remark 3.2 When actually realized for the puspose of developing approximation schemes for the PIDDS, the constructs appearing in Theorem 3.1 take the following form. The space Z is of course $R^n \times L_2^n(-r,0)$, $Z_N = R^n \times L_2^n(-r_\nu^N, 0)$, X_N is a finite dimensional subspace of Z_N such as the AVE or spline subspaces discussed in [28], I_N is the operator that takes $(\eta,\phi) \in Z$ into $\tilde{z} = (\eta, \tilde{\phi})$ in Z_N where $\tilde{\phi}$ is the restriction of ϕ to $[-r_\nu^N, 0]$ and $c(z)$ might for example be chosen from among the Padé rational function approximations to the exponential (see [28], [30]). Once a basis for X_N has been chosen, A_N can be represented by a matrix as can the operators $c\left(\frac{r_\nu^N}{N} A_N\right)^k$ $k=0,1,2\ldots\rho^N$. If the A_N are constructed and $c(z)$ is chosen so as to comply with the hypotheses and conditions of Theorem 3.1, for $z_0 \in Z$ and $t_k^N = \frac{kr_\nu^N}{N} \in [0,T]$ $k=0,1,2\ldots\rho^N$ we have that $z(t_k^N) = S(t_k^N)z_0$ is approximated by $z_k^N = c\left(\frac{r_\nu^N}{N} A_N\right)^k P_N z_0$. The construction of X_N and A_N and the selection of $c(z)$ so as to lead to convergent approximation schemes is examined in detail in Section 5.

Remark 3.3 Implicit in condition (1) in Theorem 3.1 above is the assumption that $P_N \mathcal{D} \subset D(A_N)$ $N=1,2\ldots$. However, as has been remarked above, in practice, X_N is chosen to be finite dimensional in which case $A_N : X_N \to X_N$ is a bounded operator with $D(A_N) = X_N$.

Remark 3.4 In what is to follow we shall frequently refer to $A_N \in G(M,\beta)$ as the spatial stability condition and (3.3) as the temporal stability condition.

It is not surprising that an estimate of the rate of convergence in Theorem 3.1 would depend upon both the degree to which the A_N approximate A and the degree to which $c(z)$ approximates e^z. An application of Theorem 3.2 of [7] and arguments similar to those used to verify Theorem 4.17 of [28] can be used to establish the following Theorem.

Theorem 3.2 Under the hypotheses and conditions of Theorem 3.1 suppose that $\mathcal{B}$ is a subset of $D(A^2) \cap \mathcal{D}$ for which

(1) For each $z \in \mathcal{B}$ there exists a $K=K(z)$ such that

$$|A_N P_N z - P_N A z|_N \leq K(z)/N^p$$

(2) There exists a subset $\mathcal{B}_1$ of $\mathcal{B}$ such that for $z \in \mathcal{B}_1$ and $\lambda \in C$ with $\mathrm{Re}\lambda > \beta$

(a) $S(t)z \in \mathcal{B}$ $\quad t \in [0,T]$

(b) $S(t)(\lambda I - A)z \in \mathcal{B}$ $\quad t \in [0,T]$

and the constants guaranteed by (1) for (a) and (b) are independent of $t \in [0,T]$.

Then for each $z \in \mathcal{B}_1 \cap D(A^{m+1})$ for which $(\lambda I - A)^j z \in \mathcal{B}$ $\; j = 0,1,2\ldots m$ there exist constants $K_1 = K_1(z)$ and $K_2 = K_2(z)$ such that

$$\left| c\left(\frac{r_\nu^N}{N} A_N\right)^k P_N z - P_N S\left(\frac{k r_\nu^N}{N}\right) z \right|_N \leq K_1\left(\frac{r}{N}\right)^p + K_2\left(\frac{r}{N}\right)^m$$

$k = 0,1,2\ldots\rho^N$.

4. Approximation Schemes for the PIDDS

In this section we employ the approximation framework outlined in the previous section in order to define the approximating parameter identification problems. We then go on to demonstrate that if constructed appropriately, the solutions to the approximating problems converge to a solution of the PIDDS.

Let $\pi^0 : Z \to R^n$, $\pi^1 : Z \to L_2^n(-r,0)$ be the canonical coordinate projections of Z given by $\pi^0(\eta,\phi) = \eta$ and $\pi^1(\eta,\phi) = \phi$ respectively. For $\{q_N\}$ a sequence of elements in Q with

$$q_N = (\alpha_N, h_N) = (\alpha_N, (r_1^N, r_2^N, \ldots r_\nu^N))$$

let $Z_N = Z_N(q_N) = R^n \times L_2^n(-r_\nu^N, 0)$, let $X_N = X_N(q_N)$ be a closed subspace of Z_N, let $\Pi_N = \Pi_N(q_N)$ be the orthogonal projection of Z_N onto X_N with respect to the Z_N inner product $\langle \cdot, \cdot \rangle_N$ defined by

$$\langle (\eta,\phi), (\zeta,\psi) \rangle_N = \eta^T \zeta + \int_{-r_\nu^N}^{0} \phi(\theta)^T \psi(\theta) g^N(\theta) d\theta$$

where g^N is a positive weighting function which will be described in Section 5, and let $I_N = I_N(q_N): Z \to Z_N$ be the mapping which takes $(\eta,\phi) \in Z$ into $(\eta,\tilde{\phi}) \in Z_N$ where $\tilde{\phi}$ denotes the restriction of ϕ to $[-r_\nu^N, 0]$. Define $P_N = P_N(q_N): Z \to X_N$ by $P_N(q_N) = \Pi_N(q_N) I_N(q_N)$ and let $A_N(q_N)$ be a linear transformation defined on X_N with range contained in X_N. Finally let $c(z)$ and $d(z)$ be rational functions of the complex variable z and let θ be a fixed positive scalar with $0 \le \theta \le 1$. With these definitions in hand, the approximating parameter identification problems can be stated as follows.

(NPIDDS): Given an input/output pair $(u,\zeta) \in PC^m(0,T) \times C^\ell(0,T)$ for some $T > 0$ find $\gamma_N^* = (\eta_N^*, \phi_N^*, q_N^*) = (\eta_N^*, \phi_N^*, \alpha_N^*, h_N^*) \in \Gamma$ which minimizes

$$J_N(\gamma) = |y_0^N(\gamma;u) - \zeta(0)|_{w_1}^2 + |y_{\rho_N}^N(\gamma;u) - \zeta(T)|_{w_2}^2$$

$$+ \frac{r_\nu}{N} \sum_{j=0}^{\rho^N - 1} |y_j^N(\gamma;u) - \zeta(j\frac{r_\nu}{N})|_{w_3}^2$$

subject to

(4.1) $$z_j^N = c\left(\frac{r_\nu}{N} A_N(q)\right)^j P_N(q)(\eta,\phi) +$$

$$\frac{r_\nu}{N} \sum_{\ell=1}^{j} c\left(\frac{r_\nu}{N} A_N(q)\right)^{j-\ell} d\left(\theta \frac{r_\nu}{N} A_N(q)\right) P_N(q) \hat{B}(\alpha) u\left(\frac{\ell r_\nu}{N}\right)$$

$$y_j^N = \hat{C}(\alpha)z_j^N + D(\alpha)u\left(\frac{jr_\nu}{N}\right) \qquad j = 0,1,2,\ldots,\rho^N \tag{4.2}$$

where $\hat{B}(\alpha)$, $\hat{C}(\alpha)$, $D(\alpha)$ are as they were defined in Section 2, $\alpha = (\eta,\phi,q) = (\eta,\phi,\alpha,h) = (\eta,\phi,\alpha,(r_1,r_2\ldots r_\nu))$ and ρ^N is that positive integer for which $\rho^N \frac{r_\nu}{N} \le T < (\rho^N+1)\frac{r_\nu}{N}$.

Under reasonable continuity assumptions (which will be satisfied by the specific schemes we construct in Section 5), for each N, the approximating parameter identification problem becomes the minimization of a continuous function over a compact set, and hence we are assured of the existence of a solution.

Remark 4.1 The inclusion of the operator $d(\frac{\theta r_\nu}{N} A_N(q))$ in the state equation is a consequence of the theory developed in [28]. In that paper it is shown that if $d(z)$ is chosen as a rational function approximation to the exponential for which $d(\frac{\theta r_\nu}{N} A_N(q))$ satisfies certain hypotheses (which appear in the statement of Theorem 4.1 below) then the convergence properties of the state approximation will be enhanced.

Remark 4.2 If for each $q \in Q$ and $N = 1,2\ldots$ we define the operators $B^N(q): R^m \to X_N$ and $A^N(q): X_N \to X_N$ by $B^N(q)u = d(\frac{\theta r_\nu}{N} A_N(q))P_N(q)\,\hat{B}(\alpha)u = d(\frac{\theta r_\nu}{N} A_N(q))P_N(q)(B(\alpha)u,0)$ and $A^N(q) = c(\frac{r_\nu}{N} A_N(q))$ respectively and let $z_N^0(\gamma) = P_N(q)(\eta,\phi)$ and $u_j^N = u(\frac{jr_\nu}{N})$ $j = 0,1,2\ldots\rho^N$, it is immediately clear that (4.1) is the classical variation of parameters solution to the linear non-homogeneous difference equation in X_N given by

$$z_j^N = A^N(q)z_{j-1}^N + B^N(q)u_{j-1}^N \qquad j = 1,2\ldots\rho^N \tag{4.3}$$

with initial condition $z_0^N = z_N^0(\gamma)$. Furthermore with the state equation written in the form given by (4.3) and with the exception of the fact that in its most general form the admissible initial data set is infinite dimensional, the approximating parameter identification problems are easily recognized to be in the standard form of a finite dimensional discrete linear-least squares parameter identification problem for which conventional numerical methods can be used to obtain solutions (see Chapter 4 of [29]). In practice the compact admissible initial data set S (see Section 2) is almost always finite dimensional. In fact, the set S is usually chosen to be the span of a finite collection of elements $\{\hat{\psi}_1,\hat{\psi}_2\dots,\hat{\psi}_L\}\subset Z$ over a bounded subset of R^L where the unknown parameters to be determined are the coefficients.

Theorem 4.1 Suppose $\{\gamma_N^*\} = \{(z_N^{0*},q_N^*)\} = \{(\eta_N^*,\phi_N^*,q_N^*)\} = \{(\eta_N^*,\phi_N^*,\alpha_N^*,$ $(r_1^{N*},r_2^{N*},\dots,r_\nu^{N*}))\} \subset \Gamma$ is a sequence of solutions to the problems NPIDDS and there exists $\gamma^* = (z^{0*},q^*) = (\eta^*,\phi^*,q^*) = (\eta^*,\phi^*,\alpha^*,(r_1^*,\dots,r_\nu^*)) \in \Gamma$ such that $\gamma_N^*\to\gamma^*$ in the sense that (a) $q_N^*\to q^*$ in $R^{\mu+\nu}$ and (b) $z_N^{0*}\to z^{0*}$ in Z as $N\to\infty$. Suppose further that $P_N = P_N(q_N^*):Z\to X_N(q_N^*)$, $A_N = A_N(q_N^*):X_N(q_N^*)\to X_N(q_N^*)$, $A = A(q^*):D\subset Z\to Z$, $c(z)$, $\{r_\nu^{N*}\} = \{(q_N^*)_{\mu+\nu}\}$, $\rho^{N*} = \rho^{N*}(r_\nu^{N*}) = \rho^{N*}(q_N^*)$ satisfy the conditions and hypotheses of Theorem 3.1 and that

(1) The infinite collection of operators

$$\left\{c\left(\frac{r_\nu^{N*}}{N}A_N(q_N^*)\right)^k\right\}_{k=0}^{\rho^{N*}}$$

are uniformly bounded for all N sufficiently large, and

(2) For $\theta\in[0,1]$ fixed, and each $z\in Z$ we have that for the rational function $d(z)$ the operators $d(\theta\frac{r_\nu^N}{N}A_N(q_N^*))$ exist and satisfy the condition

(4.4) $$\left| d\left(\frac{\theta r_\nu^{N*}}{N} A_N(q_N^*)\right) P_N(q_N^*)z - P_N(q_N^*)Iz \right|_N \to 0 \qquad \text{as} \qquad N \to \infty \; .$$

Then

$$|P_N(q_N^*)z(t_k^N;\gamma^*,u) - z_k^N(\gamma_N^*;u)|_N \to 0$$

as $N \to \infty$ uniformly in k, $k \in \{0,1,2\ldots\rho^{N*}\}$ where z and z_k^N are given by (2.5) and (4.1) respectively and $t_k^N = \frac{kr_\nu^{N*}}{N}$ $k = 0,1,2\ldots\rho^{N*}$.

<u>Proof</u>

The existence of the operators $c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)$ for all N sufficiently large is guaranteed by Theorem 3.1. Let M_0 be such that

$$\left| c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)^k \right|_N \leq M_0 \qquad k = 0,1,2\ldots\rho^{N*}$$

for all N sufficiently large. Then

$$|P_N(q_N^*)z(t_k^N;\gamma^*,u) - z_k^N(\gamma_N^*;u)|$$

$$= \Bigg| P_N(q_N^*)S(t_k^N;q^*)z^{0*} + P_N(q_N^*)\int_0^{t_k^N} S(t_k^N - \sigma;q^*)\hat{B}(\alpha^*)u(\sigma)d\sigma -$$

$$c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)^k P_N(q_N^*)z_N^{0*} - \frac{r_\nu^{N*}}{N} \sum_{j=1}^{k} c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)^{k-j} \cdot$$

$$d\left(\theta \frac{r_\nu^{N*}}{N} A_N(q_N^*)\right) P_N(q_N^*)\hat{B}(\alpha_N^*)u\left(\frac{jr_\nu^{N*}}{N}\right)\Bigg|_N$$

$$\leq \left| P_N(q_N^*) S(t_k^N; q^*) z^{0*} - c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)^k P_N(q_N^*) z^{0*} \right|_N$$

$$+ \left| c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)^k P_N(q_N^*) (z^{0*} - z_N^{0*}) \right|_N$$

$$+ \left| P_N(q_N^*) \int_0^{t_k^N} S(t_k^N - \sigma; q^*) \hat{B}(\alpha^*) u(\sigma) d\sigma - \right.$$

$$\left. \frac{r_\nu^{N*}}{N} \sum_{j=1}^{k} c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)^{k-j} d\left(\theta \frac{r_\nu^{N*}}{N} A_N(q_N^*)\right) P_N(q_N^*) \hat{B}(\alpha_N^*) u\left(\frac{j r_\nu^{N*}}{N}\right) \right|_N$$

$$\equiv T_1^N + T_2^N + T_3^N \ .$$

The term T_1^N tends to 0 as $N \to \infty$ uniformly in k, $k \in \{0,1,2 \ldots \rho^{N*}\}$ by Theorem 3.1 while

$$T_2^N = \left| c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)^k P_N(q_N^*) (z^{0*} - z_N^{0*}) \right|_N$$

$$\leq M_0 |z^{0*} - z_N^{0*}| \to 0$$

as $N \to \infty$ uniformly in k, $k \in \{0,1,2 \ldots \rho^{N*}\}$. We next consider the term T_3^N.

$$T_3^N \leq \left| \int_0^{t_k^N} P_N(q_N^*) S(t_k^N - \sigma; q^*) \hat{B}(\alpha^*) u(\sigma) d\sigma - \right.$$

$$\left. \int_0^{t_k^N} P_N(q_N^*) S(t_k^N - \sigma; q^*) \hat{B}(\alpha^*) u_N(\sigma) d\sigma \right|_N$$

$$+ \left| \int_0^{t_k^N} P_N(q_N^*) S(t_k^N - \sigma; q^*) \hat{B}(\alpha^*) u_N(\sigma) d\sigma \right.$$

$$- \frac{r_\nu^{N*}}{N} \sum_{j=1}^{k} P_N(q_N^*) S(t_{k-j}^N; q^*) \hat{B}(\alpha^*) u\left(\frac{j r_\nu^{N*}}{N}\right)\Bigg|_N$$

$$+ \Bigg| \frac{r_\nu^{N*}}{N} \sum_{j=1}^{k} P_N(q_N^*) S(t_{k-j}^N; q^*) \hat{B}(\alpha^*) u\left(\frac{j r_\nu^{N*}}{N}\right) -$$

$$\frac{r_\nu^{N*}}{N} \sum_{j=1}^{k} c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)^{k-j} P_N(q_N^*) \hat{B}(\alpha^*) u\left(\frac{j r_\nu^{N*}}{N}\right)\Bigg|_N$$

$$+ \Bigg| \frac{r_\nu^{N*}}{N} \sum_{j=1}^{k} c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)^{k-j} P_N(q_N^*) \hat{B}(\alpha^*) u\left(\frac{j r_\nu^{N*}}{N}\right) -$$

$$\frac{r_\nu^{N*}}{N} \sum_{j=1}^{k} c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)^{k-j} d\left(\theta \frac{r_\nu^{N*}}{N} A_N(q_N^*)\right) P_N(q_N^*) \hat{B}(\alpha^*) u\left(\frac{j r_\nu^{N*}}{N}\right)\Bigg|_N$$

$$+ \Bigg| \frac{r_\nu^{N*}}{N} \sum_{j=1}^{k} c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)^{k-j} d\left(\theta \frac{r_\nu^{N*}}{N} A_N(q_N^*)\right) P_N(q_N^*) \hat{B}(\alpha^*) u\left(\frac{j r_\nu^{N*}}{N}\right) -$$

$$\frac{r_\nu^{N*}}{N} \sum_{j=1}^{k} c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)^{k-j} d\left(\theta \frac{r_\nu^{N*}}{N} A_N(q_N^*)\right) P_N(q_N^*) \hat{B}(\alpha_N^*) u\left(\frac{j r_\nu^{N*}}{N}\right)\Bigg|_N$$

$$= \mathcal{T}_1^N + \mathcal{T}_2^N + \mathcal{T}_3^N + \mathcal{T}_4^N + \mathcal{T}_5^N$$

where $u_N \in PC^m(0,T)$ is defined by

$$u_N(\sigma) = \begin{cases} u(t_k^N) & \sigma \in [t_{k-1}^N, t_k^N) \\ & k = 1,2\ldots\rho^{N*} \\ u(T) & \sigma \in [t_{\rho^{N*}}^N, T]. \end{cases}$$

For each N sufficiently large and each $t \in [0,T]$ we define the following parameterized families of bounded linear operators with domain R^n and range in X_N. For $\eta \in R^n$ and $t \in [0,T]$, let

(i) $\hat{T}_N(t)\eta = P_N(q_N^*)S(t,q^*)(\eta,0)$

(ii) $\hat{S}_N(t)\eta = P_N(q_N^*)S(t_k^N;q^*)(\eta,0) \qquad t \in [t_k^N, t_{k+1}^N)$

$k = 0,1,2\ldots\rho^{N*}$

(iii) $\hat{c}_N(t)\eta = c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)^k P_N(q_N^*)(\eta,0) \qquad t \in [t_k^N, t_{k+1}^N)$

$k = 0,1,2\ldots\rho^{N*}$

(iv) $\hat{d}_N\eta = d\left(\theta \frac{r_\nu^{N*}}{N} A_N(q_N^*)\right) P_N(q_N^*)(\eta,0)$

(v) $\hat{I}_N\eta = P_N(q_N^*)(\eta,0)$.

Using the fact that $\{S(t,q^*):t \geq 0\}$ is a C_0 semigroup of bounded linear operators on Z and Theorem 3.1 it is not difficult to show (see Lemma 9.1 of [28]) that for each $t \in [0,T]$

(4.5) $$||\hat{T}_N(t) - \hat{S}_N(t)|| \to 0 \qquad \text{as} \quad N \to \infty$$

(4.6) $$||\hat{S}_N(t) - \hat{c}_N(t)|| \to 0 \qquad \text{as} \quad N \to \infty$$

(4.7) $$||\hat{d}_N - I_N|| \to 0 \qquad \text{as} \quad N \to \infty$$

where the norm in (4.5), (4.6) and (4.7) above is the one which is induced by the uniform operator topology on $\mathcal{B}(R^n,Z_N)$, the space of all bounded linear operators with domain R^n and range in Z_N.

We now return to the terms $\mathcal{T}_j^N$ $j = 1,2,3,4,5$ and treat each one separately and in turn. Since $u \in PC^m(0,T)$ it is therefore Riemann integrable on $[0,T]$ and hence

$$\mathcal{T}_1^N = \left| \int_0^{t_k^N} P_N(q_N^*)\, S(t_k^N - \sigma; q^*)\hat{B}(\alpha^*)(u(\sigma) - u_N(\sigma))d\sigma \right|_N$$

$$= \left| \int_0^{t_k^N} \hat{T}_N(t_k^N - \sigma) B(\alpha^*)(u(\sigma) - u_N(\sigma))d \right|_N$$

$$\leq \int_0^{t_k^N} \left|\hat{T}_N\,(t_k^N - \sigma)\right| \left|B(\alpha^*)\right| \left|u(\sigma) - u_N(\sigma)\right| d\sigma$$

$$\leq M\, e^{\beta T} \left|B(\alpha^*)\right| \left|\int_0^T \left|u(\sigma) - u_N(\sigma)\right| d\sigma \rightarrow 0 \quad \text{as} \quad N \rightarrow \infty\right.$$

uniformly in k, $k \in \{0,1,2 \ldots \rho^{N*}\}$.

Using (4.5) above we have that $\|\hat{T}_N(T - \sigma) - \hat{S}_N(T - \sigma)\|$ tends to zero for each $\sigma \in [0,T]$ as $N \rightarrow \infty$. Moreover $\|\hat{T}_N(T - \sigma) - \hat{S}_N(T - \sigma)\|$ is dominated by $g(\sigma) = 2M\, e^{\beta(T - \sigma)}$ which is integrable on $[0,T]$. Therefore, by the Lebesgue dominated convergence theorem we have

$$\mathcal{T}_2^N = \left| \int_0^{t_k^N} \hat{T}_N(t_k^N - \sigma) B(\alpha^*) u_N(\sigma) d\sigma \right. -$$

$$\left. \sum_{j=1}^{k} \int_{t_{j-1}^N}^{t_j^N} \hat{S}_N(t_k^N - \sigma) B(\alpha^*) u_N(\sigma) d\sigma \right|_N$$

$$= \left| \int_0^{t_k^N} (\hat{T}_N(t_k^N - \sigma) - \hat{S}_N(t_k^N - \sigma))B(\alpha^*)u_N(\sigma)d\sigma \right|_N$$

$$\leq \int_0^T \|\hat{T}_N(T - \sigma) - \hat{S}_N(T - \sigma)\| \, |B(\alpha^*)| \, |u_N(\sigma)|d\sigma$$

$$\leq |B(\alpha^*)| \, |u|_\infty \int_0^T \|\hat{T}_N(T - \sigma) - \hat{S}_N(T - \sigma)\| \, d\sigma \to 0$$

as $N \to \infty$ uniformly in k, $k \in \{0,1,2...\rho^{N*}\}$. Using (4.6), $g(\sigma) = Me^{\beta(T-\sigma)} + M_0$ and reasoning similar to that used above we have

$$T_3^N = \left| \frac{r_\nu^{N*}}{N} \sum_{j=1}^{k} \left(P_N(q_N^*)S(t_{k-j}^N; q^*) - c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)^{k-j} P_N(q_N^*)\right) \cdot \right.$$

$$\left. \hat{B}(\alpha^*)u\left(\frac{jr_\nu^{N*}}{N}\right)\right|_N$$

$$= \left| \sum_{j=1}^{k} \int_{t_{j-1}^N}^{t_j^N} (\hat{S}_N(t_k^N - \sigma) - \hat{c}_N(t_k^N - \sigma))B(\alpha^*)u_N(\sigma)d\sigma \right|_N$$

$$\leq \int_0^{t_k^N} \|\hat{S}_N(t_k^N - \sigma) - \hat{c}_N(t_k^N - \sigma)\| \, |B(\alpha^*)| \, |u_N(\sigma)|d\sigma$$

$$\leq |B(\alpha^*)| \, |u|_\infty \int_0^T \|\hat{S}_N(T - \sigma) - \hat{c}_N(T - \sigma)\| \, d\sigma \to 0$$

as $N \to \infty$ uniformly in k, $k \in \{0,1,2\ldots\rho^{N*}\}$. Using (4.7) we have

$$T_4^N = \left| \frac{r_\nu^{N*}}{N} \sum_{j=1}^{k} c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)^{k-j} \left(P_N(q_N^*) I - d\left(\theta \frac{r_\nu^{N*}}{N} A_N(q_N^*)\right) P_N(q_N^*) \cdot \right.\right.$$

$$\left.\left. \hat{B}(\alpha^*) u\left(\frac{j r_\nu^{N*}}{N}\right)\right|_N$$

$$\leq M_0 \frac{r_\nu^{N*}}{N} \sum_{j=1}^{k} \left| (\hat{I}_N - \hat{d}_N) B(\alpha^*) u\left(\frac{j r_\nu^{N*}}{N}\right)\right|_N$$

$$\leq M_0 \frac{r_\nu^{N*}}{N} \sum_{j=1}^{k} \| \hat{I}_N - \hat{d}_N \| \; |B(\alpha^*)| \; |u|_\infty$$

$$\leq M_0 \| \hat{I}_N - \hat{d}_N \| \; |B(\alpha^*)| \; |u|_\infty \, \rho^{N*} \frac{r_\nu^{N*}}{N}$$

$$\leq M_0 |B(\alpha^*)| \; |u|_\infty T \, \| \hat{I}_N - \hat{d}_N \| \to 0 \quad \text{as} \quad N \to \infty$$

uniformly in k, $k \in \{0,1,2\ldots\rho^{N*}\}$. Finally, recalling that B has been assumed to depend continuously upon the parameters, we have

$$T_5^N = \left| \frac{r_\nu^{N*}}{N} \sum_{j=1}^{k} c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)^{k-j} d\left(\theta \frac{r_\nu^{N*}}{N} A_N(q_N^*)\right) P_N(q_N^*) \cdot \right.$$

$$\left. (\hat{B}(\alpha^*) - \hat{B}(\alpha_N^*)) u\left(\frac{j r_\nu^{N*}}{N}\right)\right|_N$$

$$\leq \frac{r_\nu^{N*}}{N} \sum_{j=1}^{k} \left| c\left(\frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)\right|^{k-j} \left| d\left(\theta \frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)\right| \, |B(\alpha^*) - B(\alpha_N^*)| \cdot$$

$$\left| u\left(\frac{j r_\nu^{N*}}{N}\right)\right|$$

$$\leq M_0 M_1 |B(\alpha^*) - B(\alpha_N^*)| \; |u|_\infty \rho^{N*} \frac{r_\nu^{N*}}{N}$$

$$\leq M_0 M_1 |u|_\infty T |B(\alpha^*) - B(\alpha_N^*)| \to 0 \quad \text{as} \quad N \to \infty$$

uniformly in k, $k \in \{0,1,2...\rho^{N*}\}$ where M_1 is the uniform bound on the operators $d\left(\theta \frac{r_\nu^{N*}}{N} A_N(q_N^*)\right)$ guaranteed to exist by the strong convergence condition given in (4.4).

Therefore

$$T_3^N = T_1^N + T_2^N + T_3^N + T_4^N + T_5^N \to 0 \quad \text{as} \quad N \to \infty$$

uniformly in k, $k \in \{0,1,2...\rho^{N*}\}$ and the theorem is proven.

Lemma 4.1 If, under the hypotheses and conditions of Theorem 4.1 we have

$$\pi^0 P_N(q_N^*) z \to \pi^0 z \tag{4.8}$$

in R^n as $N \to \infty$ for each $z \in Z$. Then

$$|y_k^N(\gamma_N^*;u) - y(t_k^N;\gamma^*,u)| \to 0$$

as $N \to \infty$ uniformly in k, $k \in \{0,1,2...\rho^{N*}\}$ where for each $\gamma \in \Gamma$, $u \in PC^m(0,T)$, $k \in \{0,1,2...\rho^{N*}\}$ and all N sufficiently large $y(t_k^N;\gamma,u)$ is given by (2.9) and $y_k^N(\gamma;u)$ is given by (4.2).

Proof

$$|y_k^N(\gamma_N^*;u) - y(t_k^N;\gamma^*,u)|$$

$$= |\hat{C}(\alpha_N^*) z_k^N(\gamma_N^*;u) + D(\alpha_N^*) u\left(\frac{kr_\nu^{N*}}{N}\right) -$$

$$\hat{C}(\alpha^*) z(t_k^N;\gamma^*,u) - D(\alpha^*) u(t_k^N)|$$

$$\leq \; |(C(\alpha_N^*) - C(\alpha^*))\pi^0 z_k^N(\gamma_N^*;u)|$$

$$+ \; |C(\alpha^*)(\pi^0 z_k^N(\gamma_N^*;u) - \pi^0 P_N(q_N^*)z(t_k^N;\gamma^*,u))|$$

$$+ |C(\alpha^*)(\pi^0 P_N(q_N^*)z(t_k^N;\gamma^*,u) - \pi^0 z(t_k^N;\gamma^*,u))|$$

$$+ |(D(\alpha_N^*) - D(\alpha^*))u(t_k^N)|$$

$$\leq \; |C(\alpha_N^*) - C(\alpha^*)|\,|z_k^N(\gamma_N^*;u)|_N$$

$$+ |C(\alpha^*)|\,|z_k^N(\gamma_N^*;u) - P_N(q_N^*)z(t_k^N;\gamma^*,u)|_N$$

$$+ |C(\alpha^*)|\,|\pi^0 P_N(q_N^*)z(t_k^N;\gamma^*,u) - \pi^0 z(t_k^N;\gamma^*,u)|$$

$$+ \; |u|_\infty |D(\alpha_N^*) - D(\alpha^*)|$$

$$\equiv \mathcal{T}_1^N + \mathcal{T}_2^N + \mathcal{T}_3^N + \mathcal{T}_4^N \; .$$

In light of the convergence guaranteed by Theorem 4.1, it is easily verified that $\{ |z_k^N(\gamma_N^*;u)|_N \}_{k=0}^{\rho^{N*}}$ lie in a bounded subset of R which is independent of N for all N sufficiently large. Therefore using the assumptions that $\alpha_N^* \to \alpha^*$ in R^μ as $N \to \infty$ and $C(\alpha)$ and $D(\alpha)$ depend continuously upon the parameters we have $\mathcal{T}_1^N \to 0$ and $\mathcal{T}_4^N \to 0$ as $N \to \infty$ uniformly in k, $k \in \{0,1,2\ldots\rho^{N*}\}$. The term $\mathcal{T}_2^N$ tends to zero as $N \to \infty$ uniformly in k, $k \in \{0,1,2\ldots\rho^{N*}\}$ as a consequence of Theorem 4.1. Finally (4.8), the fact that the set $S \equiv \{z(t;\gamma^*,u) : t \in [0,T]\}$ is a compact subset of Z (being the continuous image of a compact subset of R) and the uniform boundedness of the operators $\pi^0 P_N(q_N^*)$ imply $\pi^0 P_N(q_N^*) \to \pi^0$ uniformly on S as $N \to \infty$ and hence $\mathcal{T}_3^N \to 0$ as $N \to \infty$ uniformly in k, $k \in \{0,1,2\ldots\rho^{N*}\}$.

We can now state and prove the major result of this paper which is that in a certain sense (which will be made precise in the statement of Theorem 4.2 below) a solution γ_N^* to the N^{th} approximating parameter identification problem is in fact an approximation of a solution γ^* of the PIDDS.

Theorem 4.2 Suppose $\{\gamma_N^*\} = \{(z_N^{0*}, q_N^*)\} \subset \Gamma$ is a sequence of solutions to the problems NPIDDS. Then there exist a $\gamma^* = (z^{0*}, q^*) \in \Gamma$ and a subsequence $\{\gamma_{N_k}^*\}$ of $\{\gamma_N^*\}$ such that $\gamma_{N_k}^* \to \gamma^*$ as $k \to \infty$ in the sense that (a) $q_{N_k}^* \to q^*$ in $R^{\mu+\nu}$ and (b) $z_{N_k}^{0*} \to z^{0*}$ in Z as $k \to \infty$. If in addition $P_N = P_N(q_N^*)$, $A_N = A_N(q_N^*)$, $A = A(q^*)$, $c(z)$, $d(z)$ and $\rho^{N*} = \rho^{N*}(q_N^*)$ satisfy the hypotheses and conditions of Theorem 4.1 and if $P_N(q_N^*)$ satisfies (4.8) then γ^* is a solution of the APIDDS (and therefore to the PIDDS as well).

Proof

Since $S \subset Z$ has been assumed compact, there exists a subsequence $\{z_{N_j}^{0*}\}$ of $\{z_N^{0*}\}$ such that $z_{N_j}^{0*} \to z^{0*} \in S$ as $j \to \infty$. Similarly $Q \subset R^{\mu+\nu}$ compact implies the existence of a subsequence $\{q_{N_{j_\ell}}^*\}$ of $\{q_{N_j}^*\}$ such that $q_{N_{j_\ell}}^* \to q^* \in Q$ as $\ell \to \infty$. Letting $\gamma^* = (z^{0*}, q^*)$ and reindexing, we obtain a subsequence $\{\gamma_{N_k}^*\}$ of $\{\gamma_N^*\}$ such that $\gamma_{N_k}^* \to \gamma^* \in \Gamma$ as $k \to \infty$.

For each $\gamma = (z^0, q) = (z^0, \alpha, (r_1, r_2 \ldots, r_\nu)) \in \Gamma$, $u \in PC^m(0,T)$, $\zeta \in C^\ell(0,T)$ and all N sufficiently large we define y^N, $\hat{y}^N$, $\zeta^N \in PC^\ell(0,T)$ by

$$y^N(\sigma) = y^N(\sigma;\gamma,u) = y(t_k^N;\gamma,u) \qquad \sigma \in I_k^N \quad k = 0,1,2\ldots\rho^N$$

$$\hat{y}^N(\sigma) = \hat{y}^N(\sigma;\gamma,u) = y_k^N(\gamma;u) \qquad \sigma \in I_k^N \quad k = 0,1,2\ldots\rho^N$$

$$\zeta^N(\sigma) = \zeta^N(\sigma;\gamma) = \zeta(t_k^N) \qquad \sigma \in I_k^N \quad k = 0,1,2\ldots\rho^N$$

where $I_k^N = I_k^N(\gamma) = [t_k^N, t_{k+1}^N)$ $t_k^N = t_k^N(\gamma) = \frac{kr_\nu}{N}$ $k = 0,1,2\ldots\rho^N$ and ρ^N is that integer for which $\rho^N \frac{r_\nu}{N} \leq T < (\rho^N+1)\frac{r_\nu}{N}$ and where $y(\cdot;\gamma,u)$ and $y_\cdot^N(\gamma;u)$ are given by (2.9) and (4.2) respectively.

If $\{\gamma_N\}$ is a sequence of elements in Γ for which $\gamma_N \to \gamma \in \Gamma$, then Lemma 4.1 implies

$$(4.9) \qquad |\hat{y}^N(\sigma;\gamma_N,u) - y^N(\sigma;\gamma,u)| \to 0$$

as $N \to \infty$ uniformly in σ for $\sigma \in [0,T]$. Furthermore the continuity of $y(\cdot,\gamma,u)$ and $\zeta(\cdot)$ and the fact that length$(I_k^N) = \frac{r_\nu^N}{N} \leq \frac{r}{N} \to 0$ as $N \to \infty$ imply

$$(4.10) \qquad |y^N(\sigma;\gamma,u) - y(\sigma;\gamma,u)| \to 0$$

and

$$|\zeta^N(\sigma;\gamma_N) - \zeta(\sigma)| \to 0 \qquad \text{as} \quad N \to \infty$$

for each $\sigma \in [0,T]$. The triangle inequality, (4.9) and (4.10) imply $\hat{y}^N(\sigma;\gamma_N,u) \to y(\sigma;\gamma,u)$ for each $\sigma \in [0,T]$ and hence by the Lebesgue Dominated Convergence Theorem we have for any $\gamma \in \Gamma$

$$J(\gamma^*) = |y(0;\gamma^*,u) - \zeta(0)|_{w_1}^2 + |y(T;\gamma^*,u) - \zeta(T)|_{w_2}^2$$

$$+ \int_0^T |y(t;\gamma^*,u) - \zeta(t)|_{w_3}^2 dt$$

$$= \lim_{k\to\infty} |\hat{y}^{N_k}(0;\gamma^*_{N_k},u) - \zeta(0)|^2_{w_1} +$$

$$\lim_{k\to\infty} |\hat{y}^{N_k}(T;\gamma^*_{N_k},u) - \zeta(T)|^2_{w_2} +$$

$$\int_0^T \lim_{k\to\infty} |\hat{y}^{N_k}(t;\gamma^*_{N_k},u) - \zeta^{N_k}(t;\gamma^*_{N_k})|^2_{w_3} dt$$

$$= \lim_{k\to\infty} |y_0^{N_k}(\gamma^*_{N_k};u) - \zeta(0)|^2_{w_1} +$$

$$\lim_{k\to\infty} |y^{N_k}_{\rho N^*_k}(\gamma^*_{N_k};u) - \zeta(T)|^2_{w_2} +$$

$$\lim_{k\to\infty} \int_0^T |\hat{y}^{N_k}(t;\gamma^*_{N_k},u) - \zeta^{N_k}(t;\gamma^*_{N_k})|^2_{w_3} dt$$

$$= \lim_{k\to\infty} |y_0^{N_k}(\gamma^*_{N_k};u) - \zeta(0)|^2_{w_1} +$$

$$\lim_{k\to\infty} |y^{N_k}_{\rho N^*_k}(\gamma^*_{N_k};u) - \zeta(T)|^2_{w_2} +$$

$$\lim_{k\to\infty} \sum_{j=0}^{\rho N^*_k - 1} \int_{\frac{j r_\nu^{N^*_k}}{N_k}}^{\frac{(j+1) r_\nu^{N^*_k}}{N_k}} |\hat{y}^{N_k}(t;\gamma^*_{N_k},u) - \zeta^{N_k}(t,\gamma^*_{N_k})|^2_{w_3} dt$$

$$= \lim_{k\to\infty} \Bigg[\left| y_0^{N_k} (\gamma^*_{N_k};u) - \zeta(0) \right|^2_{w_1} + \left| y^{N_k}_{\rho^{N^*_k}} (\gamma^*_{N_k};u) - \zeta(T) \right|^2_{w_2}$$

$$+ \frac{r_\nu^{N^*_k}}{N_k} \sum_{j=0}^{\rho^{N^*_k} - 1} \left| y_j^{N_k} (\gamma^*_{N_k};u) - \zeta\left(\frac{jr_\nu^{N_k^*}}{N_k} \right) \right|^2_{w_3} \Bigg]$$

$$= \lim_{k\to\infty} J_{N_k}(\gamma^*_{N_k})$$

$$\leq \lim_{k\to\infty} J_{N_k}(\gamma)$$

$$= \lim_{k\to\infty} \Bigg[\left| y_0^{N_k} (\gamma;u) - \zeta(0) \right|^2_{w_1} + \left| y^{N_k}_{\rho^{N_k}} (\gamma;u) - \zeta(T) \right|^2_{w_2}$$

$$+ \frac{r_\nu}{N_k} \sum_{j=0}^{\rho^{N_k}-1} \left| y_j^{N_k} (\gamma;u) - \zeta\left(\frac{jr_\nu}{N_k} \right) \right|^2_{w_3} \Bigg]$$

$$= \lim_{k\to\infty} \Bigg[\left| y_0^{N_k} (\gamma;u) - \zeta(0) \right|^2_{w_1} + \left| y^{N_k}_{\rho^{N_k}} (\gamma;u) - \zeta(T) \right|^2_{w_2}$$

$$+ \sum_{j=0}^{\rho^{N_k}-1} \int_{\frac{jr_\nu}{N_k}}^{\frac{(j+1)r_\nu}{N_k}} \left| \hat{y}^{N_k}(t;\gamma,u) - \zeta^{N_k}(t;\gamma) \right|^2_{w_3} dt \Bigg]$$

$$= \lim_{k\to\infty} \Bigg[\left| \hat{y}^{N_k}(0;\gamma,u) - \zeta(0) \right|^2_{w_1} + \left| \hat{y}^{N_k}(T;\gamma,u) - \zeta(T) \right|^2_{w_2}$$

$$+ \int_0^T \left| \hat{y}^{N_k}(t;\gamma,u) - \zeta^{N_k}(t;\gamma) \right|^2_{w_3} dt \Bigg]$$

$$= |y(0;\gamma,u) - \zeta(0)|^2_{w_1} + |y(T;\gamma,u) - \zeta(T)|^2_{w_2}$$

$$+ \int_0^T \lim_{k\to\infty} |\hat{y}^{N_k}(t;\gamma,u) - \zeta^{N_k}(t,\gamma)|^2_{w_3} dt$$

$$= |y(0;\gamma,u) - \zeta(0)|^2_{w_1} + |y(T;\gamma,u) - \zeta(T)|^2_{w_2}$$

$$+ \int_0^T |y(t;\gamma,u) - \zeta(t)|^2_{w_3} dt$$

$$= J(\gamma).$$

Thus $J(\gamma^*) \leq J(\gamma)$ for any $\gamma \in \Gamma$ and γ^* is a solution to the APIDDS.

5. Examples of Convergent Approximation Schemes for the PIDDS

In this section we construct specific examples of convergent approximation schemes for the PIDDS. That is, given a sequence $\{q_N\} \subset Q$ with $q_N \to \bar{q} \in Q$ as $N \to \infty$, for each $N = 1,2...$ we define X_N, a closed subspace of $Z_N = R^n \times L_2^n(-r_\nu^N, 0)$, $\Pi_N : Z_N \to X_N$ the orthogonal projection of Z_N onto X_N, linear operators $A_N(q_N): X_N \to X_N$ and choose rational functions $c(z)$ and $d(z)$ which satisfy the hypotheses and conditions of theorem 4.2. We require

(5.1) There exist constants M and β such that $A_N(q_N) \in G(M,\beta)$ on X_N for all N sufficiently large and $A = A(\bar{q}) \in G(M,\beta)$ on Z.

(5.2) There exists a dense subset of Z, $\mathcal{D} \subset D(A(\bar{q}))$ such that $R_\lambda(A(\bar{q}))\mathcal{D} \subset \mathcal{D}$ for each $\lambda \in C$ with $\operatorname{Re}\lambda > \beta$ and for each $z \in \mathcal{D}$ we have

$$|A_N(q_N)P_N z - P_N A(\bar{q})z|_N \to 0 \quad \text{as} \quad N \to \infty \quad \text{where} \quad P_N = \Pi_N I_N .$$

(5.3) $\pi^0 P_N z \to \pi^0 z$ for each $z \in Z$ where $\pi^0 : \begin{cases} Z \\ Z_N \end{cases} \to R^n$ is defined by $\pi^0(\eta,\phi) = \eta$.

(5.4) $c(z)$ is a rational function approximation to the exponential for which

(a) $|c(z) - e^z| = O(|z|^{m+1})$ as $z \to 0$ with $m > 0$

(b) $\deg c(z) \le m+1$

(c) $c(z)$ has no poles in $\{z \in C : \operatorname{Re} z \le 0\}$.

(d) There exists a constant M_0 such that

$$\left| c\left(\frac{r_\nu^N}{N} A_N(q_N)\right)^k \right| \le M_0 \qquad k = 0,1,2\ldots\rho^N$$

for all N sufficiently large where ρ^N is that positive integer for which $\rho^N \frac{r_\nu^N}{N} \le T < (\rho^N+1)\frac{r_\nu^N}{N}$.

(5.5) $d(z)$ is a rational function approximation to the exponential for which

(a) the operators $d\left(\frac{\theta r_\nu^N}{N} A_N(q_N)\right)$ exist for all N sufficiently large

(b) $\left| d\left(\frac{\theta r_\nu^N}{N} A_N(q_N)\right) P_N z - P_N I z \right|_N \to 0$ as $N \to \infty$ where $0 \le \theta \le 1$.

For a given choice of X_N, Π_N, $A_N(q_N)$, $c(z)$ and $d(z)$, the triple $\{X_N,\Pi_N,A_N(q_N)\}$ will be referred to as the state approximation, while the collection $\{X_N,\Pi_N,A_N(q_N),c(z),d(z)\}$ itself will be referred to as an approximation scheme. We shall consider two particular families of state approximations which can be shown to satisfy conditions (5.1), (5.2) and (5.3) above. The first, and more primitive of the two is the averaging or AVE state approximation ([4], [5], [7], [28]) in which the functional component of the subspace X_N is chosen to be the span of a finite collection of piecewise constant functions defined on $[-r_\nu^N,0]$. The second family of state approximations is spline based and is known as the SPL state approximation ([7], [9], [28]). In this case the subspace X_N is chosen to be the span of a finite collection of elements in Z_N having first or higher order spline functions as their functional component. We note that in both the AVE and SPL state approximations X_N is finite dimensional.

Once a state approximation has been set, rational functions $c(z)$ and $d(z)$ must be chosen in order to complete the construction of the approximation scheme. Although others are available, we shall restrict our attention to choices of $c(z)$ and $d(z)$ from the Padé table of rational function approximations to the exponential ([28], [30]). We shall demonstrate that for appropriate choices for $c(z)$ and $d(z)$ taken from the Padé table, the AVE and SPL state approximations generate approximation schemes which satisfy conditions (5.1) through (5.5) above and hence yield approximate solutions to the PIDDS.

All of the ideas discussed in this section have appeared elsewhere. In particular, since our discrete schemes are based upon the semi-discrete approximation schemes for the PIDDS developed by Banks, Burns and Cliff [7], the AVE and SPL state approximations are the same as those used in [7] for a similar purpose. Furthermore, since our schemes are also based upon the discrete approximation framework for the integration of LRFDE initial value problems developed in [28], the theory underlying the appropriate choice of the rational functions $c(z)$ and $d(z)$ can be found in [28]. Therefore, the construction of the state approximations, the choosing of the rational functions and the arguments used in the verification of conditions (5.1) through (5.5) for the particular schemes will only be outlined and summarized here. For a detailed explanation of the various constructs which we define and the verification of the many results which we state without proof, the interested reader is advised to consult [7] and [28].

Central to our discussion of the state approximations will be the notion of dissipativeness of an operator. A closed linear operator $\mathcal{T}:D(\mathcal{T}) \subset H \to H$ with dense domain and range in a Hilbert space H is said to be dissipative if

$$\mathrm{Re} < \mathcal{T}f, f > \leq 0$$

for each $f \in D(\mathcal{T})$. A dissipative operator is said to be maximal dissipative if it does not have a non-trivial dissipative extension. Clearly if either H is finite dimensional or $\mathcal{T}$ is bounded then if $\mathcal{T}$ is dissipative it is maximal dissipative. Standard results from linear semigroup theory ([20], [21], [23]) can be used to show that if there exists an inner product $<\cdot,\cdot>_1$ on H, equivalent to the standard inner product on H (with $m_1|\cdot| \le |\cdot|_1 \le m_2|\ |$) for which the operator $\mathcal{T} - \beta I$ is maximal dissipative for some β then $\mathcal{T}$ is the infinitesimal generator of a C_0 semigroup of bounded linear operators on H, $\{T(t): t \ge 0\}$ such that $|T(t)| \le \frac{m_2}{m_1} e^{\beta t}$. That is $\mathcal{T} \in G(M,\beta)$ on H where $M = \frac{m_2}{m_1}$.

For $q = (\alpha, r_1, r_2, \ldots r_\nu) \in Q$ we define the weighting function $g(\cdot;q)$ on $[-r,0]$ by

$$g(\theta;q) = \begin{cases} 1 & -r \le \theta < -r_{\nu-1} \\ 2 & -r_{\nu-1} \le \theta < -r_{\nu-2} \\ \cdot & \\ \cdot & \\ \cdot & \\ \cdot & \\ \cdot & \\ \cdot & \\ \nu - 1 & -r_2 \le \theta < -r_1 \\ \nu & -r_1 \le \theta \le 0 \end{cases}$$

and the inner product on Z, $\langle\cdot,\cdot\rangle_q$ by

$$(5.6) \qquad \langle(\eta,\phi), (\zeta,\psi)\rangle_q = \eta^T\zeta + \int_{-r}^{0} \phi(\theta)^T\psi(\theta)g(\theta;q)d\theta \quad .$$

It is easily seen that the q-inner product defined above is equivalent to the standard inner product on Z. In fact we have

$$(5.7) \qquad |\cdot| \le |\cdot|_q \le \sqrt{\nu}\,|\cdot| \; .$$

If we recall the definitions of $\mathcal{A}(q)$ and $L(q)$ given in section 2 using arguments similar to those found in [9] and [28] it can be shown that for $q \in Q$ and $z \in D = D(\mathcal{A}(q))$

$$(5.8) \qquad \langle A(q)z,z\rangle_q \leq \omega(q)|z|_q^2$$

where

$$(5.9) \qquad \omega(q) = \frac{\nu+1}{2} + |A_0(\alpha)| + \tfrac{1}{2}\sum_{i=1}^{\nu}|A_i(\alpha)|^2 + \tfrac{1}{2}\int_{-r_\nu}^{0}|K(\alpha,\theta)|^2 d\theta \ .$$

Since Q is a compact subset of $R^{\mu+\nu}$ and the system coefficients have been assumed to depend continuously upon the parameters we have that there exists a $\beta > 0$ such that $\omega(q) \leq \beta$ for all $q \in Q$. It is also not difficult to show (see [25]) that for $\lambda \in C$ with $\mathrm{Re}\lambda > \beta$, $R(A(q) - \lambda I) = Z$ and hence by Theorem I.4.3 of [21] we have that $A(q) - \beta I$ is a maximal dissipative operator on Z for all $q \in Q$. In light of our earlier remarks, it therefore must hold that $A(q) \in G(\sqrt{\nu},\beta)$ on Z for all $q \in Q$.

Remark 5.1 While we have defined $A(q)$ to be a mapping from $D \subset Z$ into Z for each $q \in Q$, it can also be defined as an operator from $D(A(q) \subset Z_q$ into Z_q where $Z_q = R^n \times L_2^n(-r_\nu,0)$ and

$$D(A(q)) = \{(\eta,\phi) \in Z_q \mid \eta = \phi(0),\ \phi \in W_{1,2}^n(-r_\nu,0)\} \ .$$

In both cases $A(q)(\phi(0),\phi) = (L(q)\phi, D\phi)$ where in the first case $D\phi$ is defined on $[-r,0]$ and in the second case on $[-r_\nu,0]$. In either case, however, the two definitions lead to essentially the same operator and hence we use them interchangeably.

5.1 The AVE State Approximation

Let $\{q_N\} = \{(\alpha_N, r_1^N, r_2^N, \ldots r_\nu^N)\} \subset Q$ be given with $q_N \to \bar{q} \in Q$. Define $\chi_j^N \in L_2^n(-r_\nu^N,0)$ to be the characteristic function of the interval $\left[-j\frac{r_\nu^N}{N}, -(j-1)\frac{r_\nu^N}{N}\right)$ $j = 2,3,\ldots N$ and χ_1^N to be the characteristic function of the interval $\left[\frac{-r_\nu^N}{N}, 0\right]$. Let X_N be the closed subspaces of $Z_N = R^n \times L_2^n(-r_\nu^N, 0)$ given by

$$X_N = \{ (\eta,\phi) \in Z_N \;|\eta \in R^n, \; \phi = \sum_{j=1}^{N} v_j^N \chi_j^N, \; v_j^N \in R^n\} \; .$$

With X_N as above, the orthogonal projection Π_N of Z_N onto X_N with respect to the standard innerproduct on Z_N can be computed and is given by

$$\Pi_N(\eta,\phi) = \left(\eta, \sum_{j=1}^{N} \phi_j^N \chi_j^N\right)$$

where

$$\phi_j^N = \frac{N}{r_\nu^N} \int_{\frac{-jr_\nu^N}{N}}^{\frac{-(j-1)r_\nu^N}{N}} \phi(\theta)d\theta \qquad j = 1,2,\ldots N \; .$$

In order to define the operator $A_N(q_N)$ we first define the operators $L_N(q_N):X_N \to R^n$ and $D_N(q_N):X_N \to L_2^n(-r_\nu^N,0)$ by

$$L_N(q_N)(\eta, \sum_{j=1}^{N} v_j^N \chi_j^N) = A_0(\alpha_N)\eta + \sum_{i=1}^{\nu} \sum_{j=1}^{N} A_i(\alpha_N) v_j^N \chi_j^N(-r_i^N)$$

$$+ \frac{r_\nu^N}{N} \sum_{j=1}^{N} K_j^N(\alpha_N) v_j^N$$

where $$K_j^N(\alpha) = \frac{N}{r_\nu^N} \int_{\frac{-jr_\nu^N}{N}}^{\frac{-(j-1)r_\nu^N}{N}} K(\alpha,\theta)d\theta \qquad j = 1,2,\ldots N$$

and

$$D_N(q_N)(\eta, \sum_{j=1}^{N} v_j^N \chi_j^N) = \sum_{j=1}^{N} \frac{N}{r_\nu^N} (v_{j-1}^N - v_j^N)\chi_j^N$$

where $v_0 \equiv \eta$ respectively. Let $A_N(q_N):X_N \to X_N$ be given by

$$(5.10) \qquad A_N(q_N)(\eta,\phi) = (L_N(q_N)(\eta,\phi),\ D_N(q_N)(\eta,\phi)) .$$

Let $J_N = \{j_1^N, \ldots j_{\nu-1}^N\}$ be an index set where $j_\nu^N = N$ and j_i^N is the index such that $-r_i^N \in \left[-j_i^N \frac{r_\nu^N}{N}, -(j_i^N - 1)\frac{r_\nu^N}{N}\right]$ $i = 1,2\ldots\nu-1$. Define the numbers $\{a_j^N\}_{j=1}^{N}$ by the following recurrence relation. Let $a_N^N = 1$ and

$$a_j^N = \begin{cases} a_{j+1}^N + 1 & \text{if } j \in J_N \\ & \\ a_{j+1}^N & \text{if } j \notin J_N \end{cases}$$

We next define the piecewise constant weighting function $g_N(\cdot;q_N)$ by $g_N(\theta;q_N) = a_j^N$ for $\frac{-jr_\nu^N}{N} \le \theta \le \frac{-(j-1)r_\nu^N}{N}$ $j = 1,2\ldots N$ and define the inner-product $\langle\cdot,\cdot\rangle_{g_N}$ on Z_N by

$$\langle(\eta,\phi),\ (\zeta,\psi)\rangle_{g_N} = \eta^T\zeta + \int_{-r_\nu^N}^{0} \phi(\theta)^T \psi(\theta) g_N(\theta;q_N)d\theta .$$

The $\langle\cdot,\cdot\rangle_{g_N}$ innerproduct is equivalent to the standard innerproduct on Z_N and in fact

$$(5.11) \qquad |\cdot|_N \le |\cdot|_{g_N} \le \sqrt{\nu}|\cdot|_N .$$

It can now be shown (see [7]) that for $A_N(q_N)$ as it is given in (5.10) above and $z \in X_N$

$$(5.12) \qquad \langle A_N(q_N)z,z\rangle_{g_N} \leq \omega(q_N)|z|^2_{g_N} \leq \beta|z|^2_{g_N}$$

where $\omega(q)$ is given by (5.9). Since X_N is finite dimensional and therefore $D(A_N(q_N)) = X_N$ we have $A_N(q_N) \in G(\sqrt{\nu},\beta)$ $N = 1,2\ldots$ and condition (5.1) is satisfied. In addition it can also be shown (see [25]) that

$$(5.13) \qquad \left|I + \frac{r_\nu^N}{N} A_N(q_N)\right|_{g_N} \leq 1 + K(q_N)\frac{r_\nu^N}{N} \leq 1 + \frac{\alpha r}{N}$$

where α is a constant independent of N and $q \in Q$. The bound given in (5.13) is a somewhat stronger result than dissipativeness in that (5.13) implies (5.12) with $\beta = \frac{\alpha}{2}$ (see Lemma 5.15 of [28]). The importance of condition (5.13) will become clear when the choosing of the rational function $c(z)$ is discussed in Subsection 5.3.

If we let $\mathcal{D} = \{(\phi(0),\phi) \in Z | \phi \in C_1^n(-r,0)\}$, then $\mathcal{D}$ is a dense subset of Z, $\mathcal{D} \subset D = D(A(\bar{q}))$ and for $\lambda \in C$ with $\mathrm{Re}\,\lambda > \beta$ we have $R_\lambda(A(\bar{q})\mathcal{D} \subset D(A^2(\bar{q})) \subset \mathcal{D}$. Moreover, it can be shown (see [7]) that for $z \in \mathcal{D}$

$$|A_N(q_N)P_Nz - P_NA(\bar{q})z|_N \to 0 \quad \text{as} \quad N \to \infty$$

and condition (5.2) is satisfied.

Finally, for $(\eta,\phi) \in Z$ we have

$$\pi^0 P_N z = \pi^0(\eta, \sum_{j=1}^N \phi_j^N \chi_j^N) = \eta = \pi^0 z$$

and hence we have that conditions (5.1), (5.2) and (5.3) are satisfied for AVE state approximation.

5.2 The SPL State Approximation

In this subsection we describe spline based state approximations using first order or linear splines. All of the results stated below can be modified so as to be applicable to spline based state approximations employing higher order splines.

Once again we assume $\{q_N\} = \{(\alpha_N, r_1^N, r_2^N, \dots r_\nu^N)\} \subset Q$ with $q_N \to \bar{q} \in Q$ as $N \to \infty$. We partition each of the subintervals $[-r_k^N, -r_{k-1}^N]$ $k = 1,2\dots\nu$ into N equal subintervals to define the partition $\{\theta_j^N\}_{j=1}^{\nu N}$ of $[-r_\nu^N, 0]$ where

$$\theta_j^N = -(j-(k-1)N)(r_k^N - r_{k-1}^N)/N + r_{k+1}^N$$

$j = (k-1)N, \dots kN$, $k = 1,2\dots\nu$, and define the finite dimensional subspace X_N of Z_N by

$$X_N = \left\{(\phi(0),\phi) \in Z_N \,\middle|\, \phi \text{ is a first order spline with knots at } \{\theta_j^N\}_{j=1}^{\nu N}\right\}.$$

We let Π_N be the orthogonal projection of Z_N onto X_N with respect to the $\langle\cdot,\cdot\rangle_{q_N}$ innerproduct defined in (5.6). Finally we let $A_N(q_N): X_N \to X_N$ be given by

$$A_N(q_N) = \Pi_N A(q_N) \Pi_N. \tag{5.14}$$

We note that

$$R(\Pi_N) = X_N \subset \{(\eta,\phi) \in Z_N \mid \eta = \phi(0), \phi \in W^n_{1,2}(-r_\nu^N, 0)\}$$

$$= D(A(q_N))$$

and hence the expression for $A_N(q_N)$ given by (5.14) is well defined.

Using (5.8), (5.14) and the fact that Π_N is the orthogonal projection of Z_N onto X_N with respect to the $<\cdot,\cdot>_{q_N}$ inner product we have for all $z \in X_N$

$$(5.15) \qquad \begin{aligned} <A_N(q_N)z,z>_{q_N} &= <\Pi_N A(q_N)\Pi_N z, z>_{q_N} \\ &= <A(q_N)\Pi_N z, \Pi_N z>_{q_N} \le \omega(q_N)|\Pi_N z|^2_{q_N} \\ &\le \omega(q_N)|z|^2_{q_N} \le \beta|z|^2_{q_N} \end{aligned}$$

where once again $\omega(q)$ is given by (5.9). Since X_N is finite dimensional with $D(A_N(q_N)) = X_N$ we have therefore that $A_N(q_N) \in G(\sqrt{\nu},\beta)$ $N = 1,2\ldots$ and condition (5.1) is satisfied.

Next, if we define $\mathcal{D} = D(A^3(\bar{q}))$, we have that $\mathcal{D}$ is a dense subset of Z (see [23]) and for $\lambda \in C$ with $\mathrm{Re}\lambda > \beta$, $R_\lambda(A(\bar{q}))\,\mathcal{D} \subset \mathcal{D}$. Using the properties of interpolatory splines, the fact that Π_N is an orthogonal projection (and hence has certain minimality properties) and the norm equivalence relation

$$(5.16) \qquad |\cdot|_N \le |\cdot|_{q_N} \le \sqrt{\nu}|\cdot|_N$$

given by (5.7) it can be shown that

$$|A_N(q_N)P_N z - P_N A(\bar{q})z|_N \to 0 \quad \text{as} \quad N \to \infty$$

for each $z \in \mathcal{D}$. Furthermore it can also be argued that for $\hat{\phi} \in \mathcal{D}$, $|\Pi_N I_N\hat{\phi} - I_N\hat{\phi}|_N \to 0$ as $N \to \infty$. However $\mathcal{D}$ is a dense subset of Z and the operators $\{\Pi_N I_N - I_N\}$ are uniformly bounded. Recalling that $P_N = \Pi_N I_N$ it follows therefore that $|P_N z - I_N z|_N \to 0$ as $N \to \infty$ for all $z \in Z$. This in turn implies that

$$\pi^0(P_N z) \to \pi^0 z$$

for all $z \in Z$ and the SPL state approximation defined above satisfies conditions (5.1), (5.2) and (5.3).

5.3 Selecting the Rational Functions $c(z)$ and $d(z)$

Our primary objective in this subsection is to summarize the theory developed in [28] for the selection of rational functions $c(z)$ and $d(z)$ which satisfy conditions (5.4) and (5.5) respectively for a given state approximation triple $\{X_N, \Pi_N, A_N(q_N)\}$. For a given approximation scheme $\{X_N, \Pi_N, A_N(q_N), c(z), d(z)\}$ the most difficult condition to verify is the temporal stability condition (5.4)(d). As we shall soon see, it is the happy circumstance that the relatively easily verified spatial stability condition (5.1) (which we already know is satisfied by the AVE and SPL state approximations) is, under the appropriate hypotheses, sufficient to guarantee that (5.4)(d) holds as well.

Definition 5.1 We shall say that a rational function $r(z)$ of the complex variable z is acceptable if

(1) $|r(z) - e^z| = O(|z|^{m+1})$ as $z \to 0$ $m > 0$

(2) $|r(z)| \leq 1$ $z \in \{z \in C : \mathrm{Re}\, z \leq 0\}$.

Although there are many families of rational functions which admit acceptable subclasses, among the most widely studied are the Padé rational function approximations to the exponential. The Padé approximations, which can be arranged in a tableau $\{p_{jk}(z)\}$ commonly referred to as the Padé table, are defined by the following formulae

$$p_{jk}(z) = n_{jk}(z)/d_{jk}(z) \quad j, k = 1, 2 \ldots$$

where

$$n_{jk}(z) = \sum_{\ell=0}^{k} \frac{(j+k-\ell)!k!}{(j+k)!\ell!(k-\ell)!} z^{\ell}$$

$$d_{jk}(z) = \sum_{\ell=0}^{j} \frac{(j+k-\ell)!j!}{(j+k)!\ell!(j-\ell)!} (-z)^{\ell} \quad .$$

It is easily seen that

$$\text{(5.17)} \qquad \deg p_{jk}(z) = k - j$$

and it can be shown that

$$\text{(5.18)} \qquad |p_{jk}(z) - e^z| = O(|z|^{j+k+1}) \quad z \to 0 \ .$$

Since the convergence rate estimates given in Theorem 3.2 are dependent upon the degree to which the rational function $c(z)$ approximates e^z and since the Padé approximations approximate e^z to an arbitrarily high degree, in this presentation we are content to restrict our attention to them alone. For a discussion of other families of rational function approximations to the exponential which could be employed see [28].

The following result due to Ehle [15] identifies an acceptable subclass contained in the Padé approximations.

<u>Theorem</u> 5.1 The diagonal and first two subdiagonal entries in the Padé table of rational function approximations to the exponential are acceptable. That is, the collection of rational functions given by

$$\mathscr{A}_p \equiv \{p_{n+1,n+1}(z)\}_{n=0}^{\infty} \cup \{p_{n+1,n}(z)\}_{n=0}^{\infty} \cup \{p_{n+2,n}(z)\}_{n=0}^{\infty}$$

is an acceptable subclass of the Padé approximations.

To this author's knowledge, it has not as of yet been demonstrated that Theorem 5.1 above identifies the entire acceptable subclass contained in the Padé approximations. Ehle [15], and more recently, other authors (see [28]) however, have provided evidence to the fact that this is indeed the case.

From Definition 5.1, (5.17), (5.18) and Theorem 5.1 it is immediately clear that the rational functions contained in the class $\mathscr{A}_p$ satisfy conditions (5.4)(a), (b) and (c). We next turn our attention to the temporal stability condition (5.4)(d).

The following definition and result due to J. von Neumann [26] will prove useful in our discussion below.

Definition 5.2 A set $Z \subset C$ (completed by the point at infinity) will be called a Spectral Set for the bounded linear transformation T on the Hilbert space H if (a) it is closed, (b) $Z \supseteq \sigma(t)$ and (c) for every rational function $u(z)$ satisfying the inequality $|u(z)| \leq 1$ for all $z \in Z$ we have that $|u(T)| \leq 1$.

Remark 5.2 In Definition 5.2 above, the conditions $Z \supseteq \sigma(T)$ and $|u(z)| \leq 1$ for all $z \in Z$ guarantee the existence of the operator $u(T)$.

Theorem 5.2 A necessary and sufficient condition that the halfplane $\{z \in C : \mathrm{Re}\, z \leq 0\}$ be a spectral set for the bounded linear transformation T is that T be a maximal dissipative operator on H. That is $\mathrm{Re}\, \langle Tf, f \rangle \leq 0$ for all $f \in H$.

The next lemma due to Hersh and Kato [18] provides the necessary link between the ideas of acceptability and spectral sets. The proof of this result in the form in which it is stated below can be found in [28].

<u>Lemma</u> 5.1 Suppose $\{z \in C : \text{Re } z \leq 0\}$ is a spectral set for the operator $T - \beta I$ where $\beta > 0$ and T is a bounded linear operator on a Hilbert space H. Suppose further that $r(z)$ is an acceptable rational function. Then

$$|r(hT)| \leq 1 + \beta k h$$

where k is a positive constant independent of h and T.

Lemma 5.1 can now be used to determine an appropriate choice for $c(z)$ for the AVE and SPL state approximations.

<u>Theorem</u> 5.3 Let $\{X_N, \Pi_N, A_N(q_N)\}$ be either the AVE or SPL state approximation triple. Then if $c(z) \in \mathscr{A}_p$, condition (5.4) is satisfied.

<u>Proof</u>

We need only to demonstrate that condition (5.4)(d) holds. Using respectively (5.12) and (5.15) for the AVE and SPL state approximations we have for $z \in X_N$

$$< A_N(q_N)z, z>_{g_N} \leq \beta |z|^2_{g_N}$$

and

$$<A_N(q_N)z, z>_N \leq \beta |z|^2_N .$$

In either case, therefore, we have by Theorem 5.2 that $\{z \in C : \mathrm{Re}\, z \leq 0\}$ is a spectral set for $A_N(q_N)$. Theorem 5.1 and $c(z) \in \mathscr{A}_p$ imply that $c(z)$ is acceptable. Thus, by Lemma 5.1 we have

$$\left| c\left(\frac{r_\nu^N}{N} A_N(q_N) \right) \right|_{g_N} \leq 1 + \beta k \frac{r_\nu^N}{N}$$

in the case of the AVE state approximation, and

$$\left| c\left(\frac{r_\nu^N}{N} A_N(q_N) \right) \right|_{q_N} \leq 1 + \beta k \frac{r_\nu^N}{N}$$

for the SPL state approximation. Recalling the norm equivalence relation (5.11), for the AVE state approximation we have for $z \in X_N$ and $k \in \{0,1,2\ldots\rho^N\}$

$$\left| c\left(\frac{r_\nu^N}{N} A_N(q_N) \right)^k z \right|_N \leq \left| c\left(\frac{r_\nu^N}{N} A_N(q_N) \right)^k z \right|_{g_N}$$

$$\leq \left| c\left(\frac{r_\nu^N}{N} A_N(q_N) \right) \right|_{g_N}^k |z|_{g_N} \leq \left(1 + \beta k \frac{r_\nu^N}{N} \right)^k |z|_{g_N}$$

$$\leq e^{\beta k \frac{r_\nu^N}{N} k} |z|_{g_N} \leq e^{\beta k \rho^N \frac{r_\nu^N}{N}} |z|_{g_N}$$

$$\leq \sqrt{\nu}\, e^{\beta k T} |z|_N$$

and hence

$$\left| c\left(\frac{r_\nu^N}{N} A_N(q_N) \right)^k \right|_N \leq \sqrt{\nu}\, e^{\beta k T} \equiv M_0 \quad .$$

In a similar manner, using the norm equivalence relation (5.16), the same bound can be shown to hold in the case of the SPL state approximation and the theorem is proven.

Remark 5.3 Rather than applying Lemma 5.1 to the AVE and SPL state approximations directly, we could have used it to establish a somewhat more general result. Indeed, for a given spatially stable state approximation $\{X_N, \Pi_N, A_N(q_N)\}$ and acceptable rational function $c(z)$, Lemma 5.1 implies that condition (5.4)(d) is satisfied. That is the resulting approximation schemes $\{X_N, \Pi_N, A_N(q_N), c(z), *\}$ will be temporally stable.

For the AVE and SPL state approximations and $c(z) \in \mathscr{A}_p$, Theorem 5.3 quarantees that it is possible to construct an approximation scheme satisfying conditions (5.1), (5.2), (5.3) and (5.4). However for $c(z) = r(z)/s(z) \in \mathscr{A}_p$ we must have that $\deg s(z) > 0$. This implies that it will be necessary to compute $s\left(\frac{r_\nu^N}{N} A_N(q_N)\right)^{-1}$. Furthermore, in order to increase the estimated rate of convergence, we must increase $\deg s(z)$, and hence be required to invert a relatively high degree polynomial in the operator $\frac{r_\nu^N}{N} A_N(q_N)$. This is a numerically illconditioned procedure and should, if possible, be avoided. In the case of the AVE state approximation this can be achieved. Let

$$\mathscr{S}_p \equiv \{ p_{0\ell}(z) \}_{\ell=1}^{\infty} .$$

The collection $\mathscr{S}_p$ consists of the top row of the Padé table whose entries are the Maclaurin polynomials for e^z. We note that $\mathscr{S}_p \cap \mathscr{A}_p = \phi$ and observe that $\mathscr{S}_p$ consists of precisely those rational functions in the Padé table for which no operator inverse need be calculated in the computation of $p_{jk}\left(\frac{r_\nu^N}{N} A_N(q_N)\right)$. Indeed, choosing $c(z)$ from $\mathscr{B}_p$ results in an explicit approximation scheme, where as choosing $c(z)$ from $\mathscr{A}_p$ results in an implicit approximation scheme.

It can be shown (see [28] Theorem 5.17) that if

$$\left\|I + \frac{r_\nu^N}{N} A_N(q_N)\right\| = \left\|p_{01}\left(\frac{r_\nu^N}{N} A_N(q_N)\right)\right\| \leq 1 + \alpha \frac{r_\nu^N}{N}$$

for some $\alpha > 0$ independent of N and some norm $\|\ \|$ equivalent to the standard norm on X_N (with norm equivalence constants independent of N) then there exists an $\tilde{\alpha} = \tilde{\alpha}(\ell) > 0$ such that

$$\left\|p_{0,\ell}\left(\frac{r_\nu^N}{N} A_N(q_N)\right)\right\| \leq 1 + \tilde{\alpha} \frac{r_\nu^N}{N} .$$

Therefore, in light of (5.13), arguing as we did in the proof of Theorem 5.3 the following result can be established.

<u>Theorem</u> 5.4 For the AVE state approximation and $c(z) \in \mathscr{B}_p$, condition (5.4) is satisfied.

In [28] a heuristic argument in support of choosing $d(z)$ as a rational function approximation to the exponential is given. This argument is borne out empirically in that in numerical tests, enhanced convergence properties are observed for schemes constructed with $d(z)$ chosen in this way. Therefore we want to choose $d(z)$ as a rational function approximation to the exponential for which condition (5.5) is satisfied. It is easily verified (see [28] Theorem 10.3) that if $d(z)$ is chosen to satisfy condition (5.4) it will satisfy condition (5.5) as well. For the AVE state approximation, therefore, $d(z)$ can be chosen from $\mathscr{A}_p \cup \mathscr{B}_p$, while for the SPL state approximation $d(z)$ can be chosen from $\mathscr{A}_p$. However, it is shown in [28] that for the SPL state approximation $d(z)$ can actually be chosen from $\mathscr{A}_p \cup \mathscr{B}_p$ and still satisfy condition (5.5).

Finally, we can summarize the results of this section as follows. For an approximation scheme $\{X_N, \Pi_N, A_N(q_N), c(z), d(z)\}$ constructed with the AVE state approximation and $c(z)$ and $d(z)$ chosen from $\mathscr{A}_p \cup \mathscr{B}_p$, conditions (5.1) through (5.5) will be satisfied and a sequence of solutions to the resulting sequence of approximating parameter identification problems will contain a subsequence converging to a solution of the PIDDS. A similar statement can be made for approximation schemes constructed with the SPL state approximation, $c(z)$ chosen from $\mathscr{A}_p$ and $d(z)$ chosen from $\mathscr{A}_p \cup \mathscr{B}_p$.

6. Numerical Results

In this section we discuss and analyze numerical results obtained by implementing the approximation schemes developed in the previous sections and then applying them to actual parameter identification problems in which the governing control system is a linear functional differential equation of retarded type. All of the examples which follow were run on an IBM 370/158 computer using software packages written in Fortran. We provide no information regarding storage requirements or computational efficiency in that our primary objective in performing these tests was to demonstrate the feasibility of our methods.

The approximating parameter identification problems given in Section 4 were constructed using the AVE and SPL (linear spline based) state approximations defined in Section 5, $c(z) = p_{22}(z) \in \mathscr{A}_p$, $d(z) = p_{0,2}(z) \in \mathscr{B}_p$ and $\theta = .5$. The effect of variation in the choice of $c(z)$, $d(z)$ and θ

were not tested here since this was studied in [28]. We have assumed that we have been given observational data on the interval $[0,2]$ which resulted from input $u = u_\ell \in PC^1(0,2)$ where

$$u_\ell(t) = \begin{cases} 0 & t < \ell \\ 1 & \ell \le t . \end{cases}$$

The norms $|\cdot|_{w_1}$, $|\cdot|_{w_2}$, $|\cdot|_{w_3}$ which appear in (2.1) have all been taken to be the standard Euclidean norm on R^ℓ. To obtain observational data ζ, for each example the state equation was integrated using the method of steps [16], a fourth order Runge-Kutta numerical integration scheme for ordinary differential equation initial value problems, and a preselected set of true parameter values $\gamma^* = (\eta^*, \phi^*, \alpha^*, h^*)$. We emphasize that the integration method used to obtain the observational data was completely independent of the approximation schemes being tested and hence should not have contaminated our results.

The resulting finite dimensional approximating parameter identification problems were solved using a modified version of the integration package for LRFDE initial value problems developed in [28] and the IMSL [19] routine ZXSSQ, a finite difference Levenberg-Marquardt scheme for solving the problem of minimizing the sum of squares of M non-linear functions in N unknowns. The Levenberg-Marquardt algorithm is an iterative gradient projection scheme which must be provided with an initial estimate of the unknown parameters.

Since among the principal advantages of our approximation schemes is their ability to identify the delays, it is this feature which we are most interested in testing. The examples which have been included below, therefore, all have the delays in the problem among the parameters to be identified. A discussion of the performance of the schemes on examples in which the delays need not be identified can be found in [11].

Two of the four examples which appear below have also been included in [6] where they are used to test the semi-discrete schemes developed in [7]. A comparison of the performance of the two methods (based upon the two examples below, and others not included here) reveals that they exhibit similar behavior. The similarity becomes especially apparent for the cases $N = 16$ and 32, at which point the $\frac{r_\nu^N}{N}$ time step in the totally discrete schemes becomes comparable to the 1/32 time step used in the integration of the resulting approximating ordinary differential equation in the semi-discrete schemes. In addition, as N increases, the number of observational data points, ρ^N used by the totally discrete schemes increases and becomes comparable to the 101 (N independent) data points used in the testing of the semi-discrete schemes in [6]. It is interesting to note that a reasonably good fit can be achieved using relatively few observations.

Example 6.1 (Banks, Burns, Cliff [6] Example S 2.2)

In this example we identify the time delay r in the scalar first order equation given by

$$\dot{x}(t) = .05x(t) - 4.0x(t-r) + u_1(t) \tag{6.1}$$

with initial condition

(6.2) $x(0) = 1.0 \qquad x_0(s) = 1 \qquad -r \leq s \leq 0$

and output

(6.3) $y(t) = x(t)$.

Observational data was generated by using a true parameter value of $r^* = 1$. The initial estimate of the parameter was taken to be $r^{N,0} = .6$. In Table 6.1 below, for each N and each state approximation we give the final converged value for the parameter as returned by the routine ZXSSQ as a solution to the approximating parameter identification problem.

Based upon the numerical results discussed in [28], it is not surprising to find the performance of the SPL state approximation superior to that of the AVE .

N	AVE	SPL
2	.976458	.982173
4	1.11242	.984818
8	1.08012	.984677
16	1.04227	.996628
32	1.10351	1.00126
	$r^* = 1.0$	$r^* = 1.0$

Table 6.1

Example 6.2

In this example we consider the state equation (6.1), initial data (6.2) and output (6.3) of Example 6.1

$$x(t) = .05x(t) - a_1x(t-r) + u_{.1}(t)$$

$$x(0) = 1 \qquad x_0(s) = 1 \qquad -r \le s \le 0$$

$$y(t) = x(t)$$

and identify the coefficient a_1 of the delay term and the delay r itself. The true values of the parameters were taken to be $a_1^* = 4.0$ and $r^* = 1$ respectively with start up values given by $a_1^{N,0} = 3.0$ and $r^{N,0} = .6$. Our results are summarized in Table 6.2.

N	AVE		SPL	
2	Did Not Converge		Did Not Converge	
4	4.59759	1.20779	4.13681	.991267
8	Did Not Converge		4.09309	.987206
16	4.17380	1.04557	4.02157	.996570
32	4.06641	1.02561	3.99287	1.00124
	$a_1^* = 4.0$	$r^* = 1.0$	$a_1^* = 4.0$	$r^* = 1.0$

Table 6.2

Example 6.3 (Banks, Burns, Cliff [6] Example 01.2)

In this example we identify the time delay r in the damped harmonic oscillator with delayed damping and delayed restoring force given by

(6.4) $$\ddot{x}(t) + 36\,x(t) + 2.5\,\dot{x}(t-r) + 9.0\,x(t-r) = u_{.1}(t)$$

together with initial conditions and output given by

(6.5) $$x(0) = 1 \qquad \dot{x}(0) = 0$$

(6.6) $$x_0(s) = 1 \qquad \dot{x}_0(s) = 0 \qquad -r \leq s \leq 0$$

and

(6.7) $$y(t) = x(t)$$

respectively. The initial value problem (6.4), (6.5), (6.6), (6.7) can be wirtten as an equivalent first order system:

$$\dot{X}(t) = \begin{bmatrix} 0 & 1 \\ -36 & 0 \end{bmatrix} X(t) + \begin{bmatrix} 0 & 0 \\ -9.0 & -2.5 \end{bmatrix} X(t-r) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_{.1}(t)$$

$$X(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad X_0(s) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad -r \leq s \leq 0$$

$$y(t) = [1,0]\, X(t).$$

where $X(t) = \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix}$.

The true parameter value was taken to be $r^* = 1.0$ with start up value given by $r^{N,0} = 1.2$. Our results for this example, which are given in Table 6.3 once again exhibit the fact that the SPL schemes are superior to the AVE.

N	AVE	SPL
2	Did Not Converge	1.05621
4	1.22407	1.18990
8	1.14306	.991904
16	1.03183	.998599
	$r^* = 1.0$	$r^* = 1.0$

Table 6.3

Example 6.4

Here we once again consider the state equation (6.4), initial conditions (6.5), (6.6) and output (6.7) and identify the coefficient of the restoring force term and the time delay. Written as an equivalent first order system, the state equation, initial conditions and output are given by

$$\dot{X}(t) = \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix} X(t) + \begin{bmatrix} 0 & 0 \\ -9.0 & -2.5 \end{bmatrix} X(t-r) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_{.1}(t)$$

$$X(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad X_0(s) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad -r \le s \le 0$$

$$y(t) = [1,0]\, X(t).$$

respectively where $X(t) = \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix}$.

The true parameter values were taken to be $\omega^* = 6.0$ and $r^* = 1.0$ with start up values given by $\omega^{N,0} = 5.0$ and $r^{N,0} = 1.2$ respectively. Our results for this example are summarized in Table 6.4.

N	AVE		SPL	
2	4.53647	1.16643	6.26975	1.00952
4	6.28624	.895982	6.34399	.921017
8	Did Not Converge		6.05748	.985784
16	6.07952	1.04665	6.01031	.997449
	$\omega^* = 6.0$	$r^* = 1.0$	$\omega^* = 6.0$	$r^* = 1.0$

Table 6.4

In this example, as was the case in all multi-parameter, higher dimensional examples we studied, the SPL schemes performed far better than the AVE. In fact, even for large values of N, it was not uncommon for the SPL schemes to converge while the AVE schemes did not. In all examples studied, for N sufficiently large, the SPL based schemes would always produce a solution to the approximating parameter identification problem. Moreover, as N increased, the solutions to the approximating problems appeared to be converging to the true parameter values used to generate the observational data.

References

[1] H. T. Banks, Approximation of nonlinear functional differential equation control systems, J. Optimization Theory Appl., 29 (1979), 383 - 408.

[2] H. T. Banks, Identification of nonlinear delay systems using spline methods, Proc. International Conference on Nonlinear Phenomena in the Math Sciences, Univ. Texas, Arlington, June 16-20, 1980, Academic Press, to appear.

[3] H. T. Banks, Parameter identification techniques for physiological control systems, Lectures in Applied Math, 19, Amer. Math. Society, Providence, R.I., 1981.

[4] H. T. Banks and J. A. Burns, An abstract framework for approximate solutions to optimal control problems governed by hereditary sytems, Intl. Conference on Differential Equations, (H. Antosiewicz, ed.), Academic Press (1975), 10 - 25.

[5] H. T. Banks and J. A. Burns, Hereditary control problems: numerical methods based on averaging approximations, SIAM J. Control Optimization, 16 (1978), 169 - 208.

[6] H. T. Banks, J. A. Burns and E. M. Cliff, A comparison of numerical methods for identification and optimization problems involving control systems with delays, Brown Univ. LCDS Tech. Rep. 79-7 (1979), Providence, R.I.

[7] H. T. Banks, J. A. Burns and E. M. Cliff, Parameter estimation and identification for systems with delays, SIAM J. Control Optimization, to appear.

[8] H. T. Banks and P. L. Daniel, Estimation of delays and other parameters in nonlinear functional differential equations, to appear.

[9] H. T. Banks and F. Kappel, Spline approximations for functional differential equations, J. Differential Equations, 34 (1979), 496 - 522.

[10] H. T. Banks and K. Kunisch, An approximation theory for non-linear partial differential equations with applications to identification and control, Institute for Computer Applications in Science and Engineering Report 81-16 (1981), Hampton, VA.

[11] H. T. Banks and I. G. Rosen, Approximation techniques for parameter estimation in hereditary control systems, Proc. IEEE - Conference on Decision and Control (Alberquerque, Dec., 1980), 741 - 743.

[12] J. A. Burns and E. M. Cliff, Methods for approximating solutions to linear hereditary quadratic optimal control problems, IEEE Trans. Automatic Control 23 (1978), 21 - 36.

[13] J. A. Burns and P. D. Hirsch, A difference equation approach to parameter estimation for differential-delay equations, Applied Mathematics and Computation, 7 (1980), 281 - 311.

[14] P. L. Daniel, Spline-based approximation methods for the identification and control of nonlinear functional differential equations, Ph.D. Thesis, Brown Univ., Providence, R.I., June, 1981.

[15] R. L. Ehle, A-stable methods and Padé approximations to the exponential, SIAM J. Math. Anal., 4 (1973), 671 - 680.

[16] L. E. El'sgol'ts, Introduction to the Theory of Differential Equations with Deviating Arguments, Holden-Day, San Francisco, 1966.

[17] J. S. Gibson, Linear quadratic optimal control of hereditary differential systems: infinite dimensional Riccati equations and numerical approximations, preprint, (Sept., 1980), submitted to SIAM J. Control Optimization.

[18] R. Hersh and T. Kato, High-accuracy stable difference schemes for well-posed initial-value problems, SIAM J. Numer. Anal., 16 (1979), 670 - 682.

[19] International Mathematical and Statistical Libraries, Library 1, Edition 6, (1977), ZXSSQ-1-6.

[20] T. Kato, Perturbation Theory for Linear Operators, second edition, Springer-Verlag, New York, 1976.

[21] S. G. Krein, Linear differential equations in Banach space, Trans. Math. Monographs, 29, American Math. Soc., Providence, 1971.

[22] K. Kunisch, Approximation schemes for the linear quadratic optimal control problem associated with delay equations, SIAM J. Control Optimization, to appear.

[23] A. Pazy, Semigroups of linear operators and applications to partial differential equations, Math. Dept. Lecture Notes, 10, Univ. of Maryland, College Park, 1974.

[24] D. Reber, Approximation and optimal control of linear hereditary systems, Ph.D. Thesis, Brown Univ., November, 1977.

[25] D. C. Reber, A finite difference technique for solving optimization problems governed by linear functional differential equations, J. Differential Equations, 32 (1979), 192 - 232.

[26] F. Reisz and B. Sz.-Nagy, Functional Analysis, Ungar, New York, 1975.

[27] I. G. Rosen, A discrete approximation framework for hereditary systems, Ph.D. Thesis, Brown Univ., Providence, R.I., June 1980.

[28] I. G. Rosen, A discrete approximation framework for hereditary systems, J. Differential Equations, 40 (1981), 377 - 449.

[29] A. P. Sage and J. L. Melsa, System Identification, Academic Press, New York, 1971.

[30] R. S. Varga, Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, N.J., 1962, Chapter 8.3.